Volatility Jump Regressions*

Robert Davies,†

September 2, 2016

Abstract

This paper develops econometric tools for studying the jump dependencies between the underlying or latent spot volatilities of two assets from high-frequency observations on a fixed time interval – with a particular interest in the relationship between the individual volatilities of traded assets and the volatilities of aggregate risk factors such as the market volatility. The paper derives an asymptotically valid test for the stability of a linear volatility jump relationship between these assets and proposes an asymptoticly valid and consistent likelihood based estimator for the beta in such relationships. In addition, the paper proposes a bootstrap procedure for conducting inference together with a justification of its asymptotic validity.

Keywords: bootstrap, high-frequency data, jumps, regression, semimartingale, specification test, stochastic volatility.

JEL classification: C51, C52, G12.

*I would like to especially thank George Tauchen and Jia Li for their invaluable help in producing this paper. In addition, I would like to thank Tim Bollerslev, Federico Bugni, Andrew Patton, and Brian Weller as well as all the participants of Duke's financial econometrics work-group and the discussants and participants at other seminars and conferences for their helpful discussions and suggestions.

†Department of Economics, Duke University, Durham, NC 27708; e-mail: robert.davies@duke.edu.

1 Introduction

The current paper considers the relationship between the volatility of individual assets and the volatility of an aggregate risk factor, such as the market volatility, focusing specifically at the jump times in the volatility of the risk factor. If the volatility of the risk factor is itself a priced source of risk then an exploration of the link between the volatilities of individual assets and the volatility of aggregate risk factors will prove useful and insightful for the pricing of individual assets and for the understanding of the cross-section of asset returns.

To think about such a relationship let $(Y_t)_{t\geq 1}$ be the price process of some asset and $(Z_t)_{t\geq 1}$ an aggregate risk factor, such as the market, and denote their co-volatility matrix as c_t with $c_{YY,t}$ and $c_{ZZ,t}$ their respectively volatilities. The current paper studies relationships between elements of c_t at jump times $\tau \in \mathcal{T}$ where we denote the jumps in c_t as $\Delta c_t = c_t - c_{t-}$. The statistical inference is based on discrete observations of (Y, Z) sampled on an observation grid with asymptotically shrinking mesh. Because of this and the latent nature of the spot volatilities we will only be able to observe estimates of c_t. The primary focus of the paper is on generalized "volatility jump betas" which the paper defines as

$$\beta_\tau \equiv \frac{\Delta g(c_\tau)}{\Delta h(c_\tau)} \tag{1}$$

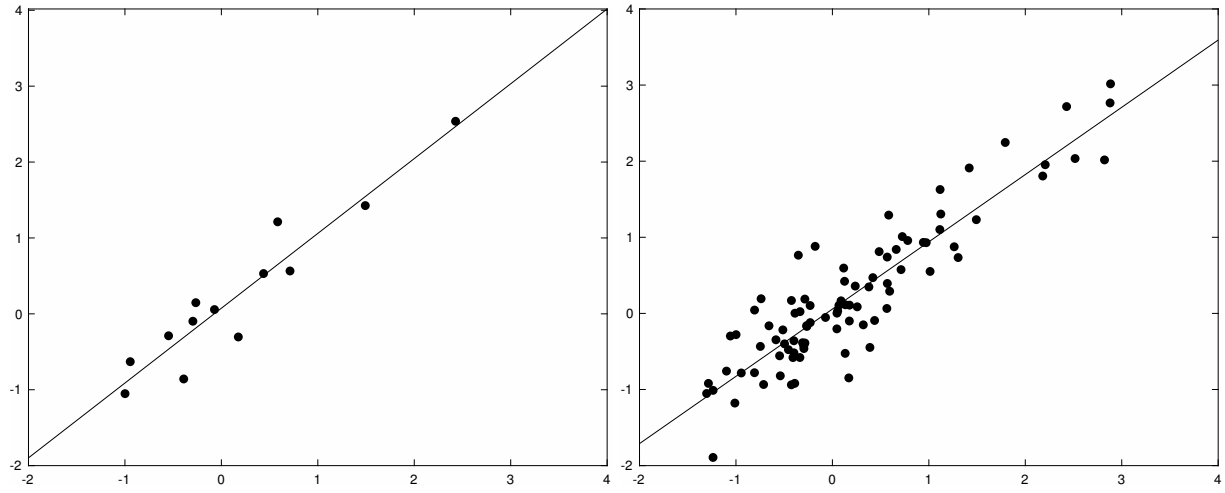
for $\tau \in \mathcal{T}$ where $g, h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ are any continuously differentiable functions and $\Delta g(c_t) = g(c_t) - g(t_{t-})$ and $\Delta h(c_t) = h(c_t) - h(t_{t-})$. Taking $g(c) = c_{YY,t}$ and $h(c) = c_{ZZ,t}$, for example, would uncover the spot-volatility beta in variance units whereas $g(c) = \log c_{YY,t}$ and $h(c) = \log c_{ZZ,t}$ or $g(c) = \sqrt{c_{YY,t}}$ and $h(c) = \sqrt{c_{ZZ,t}}$ would allow us to explore the relationship in log or standard-deviation units. We might also consider $g(c) = c_{YY,t} - (c_{ZY,t})^2/c_{ZZ,t}$ and $h(c) = c_{ZZ,t}$ which would uncover the beta between the spot volatility in Z and the "idiosyncratic" spot volatility in Y.

Without any model restriction, the spot volatility jump beta is stochastic and varies across instances of jump events. However, in many asset pricing models a linear relationship between functions of the spot volatilities might exist which would motivate testing a linear relationship of the form

$$\Delta g(c_\tau) = \beta \Delta h(c_\tau) \tag{2}$$

for $\tau \in \mathcal{T}$ where β is constant. For example, when $g(c) = \log c_{YY,t}$ and $h(c) = \log c_{ZZ,t}$ the beta in (2) is the elasticity between the spot volatilities of Z and Y. The magnitude of that beta then would tell us whether a concave, convex, or linear relationship existed between the spot volatilities

Figure 1: A Representative Illustration of Volatility Jump Regressions



Note: These figures plot estimates of the contemporaneous log-volatility co-jumps between the E-mini S&P500 index futures (ES) and the Consumer Discretionary ETF (XLY) together with linear fits. The estimated volatility jumps in the E-mini are on the horizontal axis while the those for XLY are along the vertical axis. The volatility co-jumps are estimated using returns sampled at a one-minute frequency for 2008 (left) and 2007-2015 (right). The procedure for estimating the volatility co-jumps and their jump times is detailed in Section 3.

which could potentially have very interesting results for any model of asset volatilities. Note that, equation (2) is akin to a linear regression between functions of the spot volatilities where the error in the observed equation $\Delta g(\hat{c}_\tau) = \beta \Delta h(\hat{c}_\tau) + \epsilon_\tau$ is due to the estimation error in estimating c_t. This will be the basis for the asymptotic results to follow in the paper.

A motivating empirical example of volatility jump regressions is given in Figure 1. The figure plots the estimated co-jumps in the log-volatility of the E-mini S&P500 index futures (ES), which is used here to proxy for the market, against those for the Consumer Discretionary portfolio ETF (XLY) together with a linear fit based on the model in (2). Note that even if the linear model in (2) held for all the estimated jump times that we would not expect to see a perfectly linear relationship in the plots due to the estimation error in estimating the spot volatilities. Considering this caveat the plots do show a remarkably good fit for a linear relationship in the one-year subsample (left) and a still reasonable fit for the entire nine-year sample (right). Note as well how closely the line of best fit is to $\beta = 1$ in both panels – this would imply a linear relationship between the spot volatility jumps of the E-mini and XLY.[1] The main contribution of this paper is to develop econometric techniques to both test the fit of these linear relationships and to conduct inference on the beta in such relationships.

[1]For 2008 the beta is estimated at 1.01 with a 95% confidence interval of $[0.89, 1.21]$ and for the entire 2007-2015 sample it is estimated at 0.91 with a 95% confidence interval of $[0.77, 1.00]$

The motivations for studying spot volatility jumps are several fold, but one particular motivation is the evidence that the volatilities of aggregate risk factors, or even the volatilities of individual assets, are themselves priced sources of risk. Numerous papers in the options pricing literature have documented a negative price of risk for the volatility of their underlying assets as well as the market volatility. In the options pricing context the "vega" of an option, i.e., the derivative of the option's value with respect to the volatility of the underlying asset, is widely studied both by researchers and practitioners. In terms of the cross-section of returns, one notable paper Ang, Hodrick, Xing, and Zhang (2006) explored the cross-sectional pricing of volatility risk in a fashion analogous to Fama and French (1992) and found that even controlling for the three Fama and French (1993) factors that shocks to aggregate volatility carry a statistically significant negative price of risk of approximately -1% per annum. Other papers have examined the "variance risk premium" which is defined as the difference between the expectation of future market volatility under the risk-neutral measure and its exception under the physical measure. One paper, Bollerslev, Tauchen, and Zhou (2009), found that the variance risk premium explains a considerable fraction of the variation in quarterly stock market returns. This paper explores the pricing of volatility jumps in an empirical application in Section 6.4 where, in a series of predictive regressions, future market excess returns are regressed on estimated market spot volatility jumps. The empirical applications finds spot volatility jumps to be most predictive of future excess returns at a three to four month horizon with R^2s of nearly ten percent.

In terms of economic theory the last decade has seen a considerable body of work developed that attempts to explain the underlying economic mechanisms that would lead to a negative price of volatility or a variance risk premium. An early paper, Tauchen (2011), widely circulated as a working paper before being published, showed how Epstein-Zin preferences and a model with time-varying volatility could endogenously give rise to a "leverage effect" and a variance risk premium with the sign of these effects depending critically on the coefficient of risk aversion and the intertemporal elasticity of substitution. Drechsler and Yaron (2011) show how a preference for the early resolution of uncertainty will lead to a variance risk premium in the long-run risk framework of Bansal and Yaron (2004). The intuition for their result is that if investors have a preference for the early resolution of uncertainty then they will pay to hedge against increases in uncertainty. Since volatility arises as a time-varying source of economic uncertainty in the long-run risk framework volatility itself becomes a priced risk factor. Two other recent papers to develop a variance risk premium in a long-run risk framework are Bollerslev, Sizova, and Tauchen (2011) and Bansal, Kiku, Shaliastovich, and Yaron (2014). Volatility regressions might aid in developing tests of the

implications of these theories.

In addition to the leverage effect more recent research has documented a negative relation between the idiosyncratic volatility of an asset and its subsequent returns. This has come to be known as the "idiosyncratic volatility puzzle." This phenomenon was first documented in Ang, Hodrick, Xing, and Zhang (2006). They conjectured that this finding might result because stocks with high idiosyncratic volatilities may have a high exposure to the aggregate volatility risk, which lowers their average returns. Empirically, however, they find this to be an incompletely explanation. This paper explores such a relationship in an empirical application in Section 6.3 where the beta between the idiosyncratic spot volatility jumps of various assets and the market spot volatility jumps are estimated. Near all of the betas are all found to statistically significantly different from zero implying a correlation between the idiosyncratic volatilities of assets and the market volatility at jump times. Even further, the R^2s are often in excess of 0.50 implying that the market volatility jumps often explains a substantial fraction of the variation in the idiosyncratic volatility jumps themselves. Other explanations for the idiosyncratic volatility puzzle have posited lottery-like preference on the part of some investors or have examined market frictions. A thorough review of the literature since then and these and other proposed explanations for the "puzzle" can be found in Hou and Loh (2016).

The motivation for examining the relationship between spot volatilities at jump times, particularly jump times in the market volatility, is several fold. First, there is a strong correlation between jumps in the price-level of the market and jumps in the market volatility. This is documented in Todorov and Tauchen (2011) and elsewhere. Given the strong negative correlation between returns and their volatility,[2] and the fact that jumps in the price level provide an undiversifiable source of risk in addition to an economically significant proportion of the quadratic variation in returns, examining the link between volatilities at jump times seems especially warranted. Second, many of these jump times correspond to scheduled macroeconomic announcements such as the FOMC announcements of the Federal Reserve. Understanding the link between these announcements and changes in volatilities could elucidate another channel whereby macroeconomic announcements affect financial markets. Third, given the large evidence on the time-varying nature of volatilities an exploration of volatility co-movements at jump times might illuminate an economic relationship different from that at non-jump times. On the level of returns, Li, Todorov, and Tauchen (2015) explore jump betas between asset returns and market returns at jump times. They find that for many assets that these betas remain constant at jump intervals, within a given time span, in direct

[2]Andersen, Bondarenko, and Gonzalez-Perez (2015) provide evidence that this relationship exists and remains strongly negative at both jump and non-jump intervals.

contrast to findings in Reiss, Todorov, and Tauchen (2015) and elsewhere in which such betas are found to be time varying. A similar finding might arise when examining the relationship between volatilities. Finally, due to the fact that volatilities are latent processes and must be estimated there is likely to be a fairly low signal-to-noise ratio in nearly any study of volatilities. Especially if one considers microstructure issues. A study focusing on the large jumps in volatilities could mitigate these problem because the focus would be directed only on relatively large movements in these processes where a higher signal-to-noise ratio is likely to exist and where the magnitude of changes in the true latent volatilities are likely to be considerably larger than the magnitudes of changes in the volatilities any possible microstructure factors.

Moving to the results of the paper, the main contributions can be summarized as follows. In the first part of the analysis, the paper develops a specification test for whether a linear relationship such as (2) exists between general functions of the spot co-volatility of two assets at jump times. The test is asymptotically consistent against all nonparametric fixed alternatives for which a linear relationship is violated, for example, due to time variation in the beta of the relationship and/or nonlinearities between the volatility jumps (i.e., the dependence of the volatility jump beta on the volatility jump size). Next, the paper develops tools for estimating and conducting inference on the volatility jump beta of the linear relationship. The paper show how an ordinary least-squares estimator can lead to potentially biased estimates of the volatility jump beta due to the estimation error in estimating the volatility jumps and provides a quasi-maximum-likelihood estimator based on an asymptotic approximation of the distribution of the errors in estimating the spot volatilities that corrects for this bias in small samples and asymptotically. Finally, since the asymptotic distribution of the likelihood based estimator of the volatility jump beta is highly non-standard, the paper introduces an intuitive and easy to implement bootstrap procedure for making inference and justifies its asymptotic validity.

In terms of its empirical contributions this paper contains three empirical applications of the methods and procedures introduced. The first empirical application studies the volatility jump beta between the market and a panel of assets consisting of the nine industry portfolios that comprise the S&P500 index and the thirty stocks within the Dow-Jones Industrial Average (DOW30) over the period 2007-2015, with the E-mini being the proxy for the market. Comparing log-volatility jumps the study documents that the market volatility jump beta of many financial assets appears to remain stable over a period of one year, but finds evidence for temporal variation over the entire nine year sample. Interestingly, the study finds the volatility jump beta for many assets to be very near one which would imply a linear relationship between the spot volatility of the assets and the market

volatility at jump times. The next empirical application estimates the jump beta between the idiosyncratic volatility of the DOW30 stocks and the market spot volatility. The study finds many of the betas to be statistically different from zero implying a correlation between the idiosyncratic volatilities of these assets and the market volatility at jump times. Even further, the R^2s are often in excess of 0.50 implying that the market volatility jumps often explains a substantial fraction of the variation in the idiosyncratic volatility jumps themselves. The final empirical application examines the pricing of market volatility jumps. Through a series of predictive regressions between future market excess returns and the estimated market spot volatility jumps, the study documents spot volatility jumps to be most predictive of future excess returns at a three to four month horizon with R^2s of nearly ten percent.

The rest of the paper is organized as follows. Section 2 introduces the main setting and assumptions that will be used throughout the paper. Section 3 describes the econometric framework and asymptotic theory used in estimating the spot volatilities and their jump times. Section 4 introduces the framework for estimating and evaluating volatility jump regressions and introduces a bootstrap procedure for conducting inference. Section 5 contains a Monte Carlo evaluation. Section 6 shows the results of three empirical applications. Section 7 concludes. All proofs are given in the appendix which is Section 8.

2 Setting

This section introduces the formal setup for the analysis of this paper. The following notations are used throughout. The transpose of a matrix A is denoted by $A^\top$. For two vectors a and b, the notation $a \leq b$ is meant to convey that the inequality holds component-wise. The functions $vec(\cdot)$, $\det(\cdot)$ and $\mathrm{Tr}(\cdot)$ denote matrix vectorization, determinant and trace, respectively. The Euclidean norm of a linear space is denoted $\|\cdot\|$. The notation $\mathbb{R}_*$ is used to denote the set of nonzero real numbers, that is, $\mathbb{R}_* \equiv \mathbb{R} \setminus \{0\}$. The cardinality of a (possibly random) set $\mathcal{P}$ is denoted $|\mathcal{P}|$. For any random variable ξ, the paper follows the standard shorthand notation $\{\xi$ satisfies some property$\}$ for $\{\omega \in \Omega : \xi(\omega)$ satisfies some property$\}$. The largest smaller integer function is denoted by $\lfloor\cdot\rfloor$. For two sequences of positive real numbers a_n and b_n the notation $a_n \asymp b_n$ implies that $b_n/c \leq a_n \leq cb_n$ for some constant $c \geq 1$ and all n. All limits are for $n \to \infty$. The notation $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{L}}$ and $\xrightarrow{\mathcal{L}\text{-}s}$ denote convergence in probability, convergence in law, and stable convergence in law, respectively.

2.1 The underlying processes

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. Throughout the paper, all processes are assumed to be càdlàg adapted. First, consider a 2-dimensional Itô semimartingale $X = (Y, Z)$ with the form

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta \mathbb{1}_{\{||\delta||\leq 1\}}) * (\underline{p} - \underline{q})_t + (\delta \mathbb{1}_{\{||\delta||>1\}}) * \underline{p}_t, \tag{3}$$

where the 2-dimensional process b is the drift, the 2×2-dimensional process σ is the stochastic (co)volatility process, W is a 2-dimensional Brownian motion, and $\underline{p}$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\underline{q}(dt, dx) = dt \otimes \lambda(dx)$. This model is commonly used for analyzing high-frequency data (see, e.g., Section 2.1.4 of Jacod and Protter (2012) and Section 1.4.3 of Aït-Sahalia and Jacod (2014)). It is assumed that X is observed at discrete times $i\Delta_n$, for $0 \leq i \leq n \equiv \lfloor T/\Delta_n \rfloor$, within the *fixed* time interval $[0, T]$. The increments of X are denoted by

$$\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}, \quad i = 1, \ldots, n. \tag{4}$$

The paper considers an infill asymptotic setting below in which $\Delta_n \to 0$ as $n \to \infty$.

The spot covariance matrix of X is denoted by

$$c_t = \begin{pmatrix} c_{ZZ,t} & c_{ZY,t} \\ c_{YZ,t} & c_{ZZ,t} \end{pmatrix} = \sigma_t \sigma_t^\top.$$

The spot covariance matrix process c quantifies the diffusive risk of assets, which itself evolves in a stochastic manner. In particular, during major market events (e.g., public announcements or natural disasters), various components of c tend to jump simultaneously.

In this paper, the primary interest is in the dependency among the jumps of the spot covariance matrix. More precisely, for real-valued transforms $g(\cdot)$ and $h(\cdot)$, the paper studies the linear dependency between the jumps of $g(c_t)$ and $h(c_t)$. For example,

$$\Delta g\left(c_t\right) = g\left(c_t\right) - g\left(c_{t-}\right).$$

In this paper, the aim is to develop econometric tools for studying the dependencies among the jump components of the spot covariance process.

2.2 Regularity conditions

Our basic regularity condition for X is given by the following assumption.

Assumption 1. *(a) The process b is locally bounded; (b) σ_t is nonsingular, (c) δ is a 2-dimensional predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ which satisfies $|\delta(\omega, t, z)| \leq J_n(z)$ when $t \leq \tau_n(\omega)$, where (τ_n) is a sequence of stopping times increasing to ∞ and each J_n is a non-random function satisfying $\int (J_n(z) \wedge 1)\lambda(dz) < \infty$.*

Assumption 1 is not much stronger than the property of being an Itô semimartingale, in the sense that, virtually all models using Itô semimartingales satisfy the assumption. Next we need a representation for the latent spot volatility process. As before consider the "spot volatility" matrix $c = \sigma\sigma^\top$ of X. Let

$$c_t = c_0 + \int_0^t b_s^{(c)} ds + \int_0^t \sigma_s^{(c)} dW_s^{(c)} + \delta^c * (\underline{p}^c - \underline{q}^c)_t \tag{5}$$

where $b^{(c)}$ and $\sigma^{(c)}$ are 2×2-dimensional and $2 \times 2 \times 2$-dimensional stochastic processes respectively, $W^{(c)}$ is a 2-dimensional Brownian motion, and $\underline{p}^{(c)}$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\underline{q}^{(c)}(dt, dx) = dt \otimes \lambda^{(c)}(dx)$. As before, the jump of c at time t is denoted by $\Delta c_t \equiv c_t - c_{t-}$, where $c_{t-} \equiv \lim_{s \uparrow t} c_s$. The basic regularity condition for c is given by the following assumption.

Assumption 2. *(a) The processes $b^{(c)}$ is locally bounded, (b) $c^{(c)}$ is nonsingular, (c) $\delta^{(c)}$ is a 2-dimensional predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$, and (d) $\underline{q}^{(c)}([0, T], \mathbb{R}) < \infty$.*

The representations in (3) and (5) and Assumptions 1 and 2 differ in the specification for the jumps of X and c respectively. The representation in (3) allows X to have an infinite activity jump component with only a few minor restrictions on the jump sizes and their activity given by Assumption 1(c). In contrast, the restriction imposed by Assumption 2(d) restricts c to have a finite activity jump process. This difference is mainly for simplicity as the focus of the paper is on "big" jumps is c, i.e., jumps that not "sufficiently" close to zero. This assumption could be dropped and a similar representation as in (3) could be made for c with the focus being on jumps with sizes bounded away from zero.[3] For further details on the representations in (3) or (5) or Assumptions 1 and 2 see Section 2.1.4 of Jacod and Protter (2012), Section 1.4.3 of Aït-Sahalia and Jacod (2014), or Grigelionis (1971).

3 Estimating the Latent Spot Volatility

This section details the econometric techniques used to estimate the latent spot volatilities as well as the jump times in the market volatility.

[3] Alternatively, c can be defined as "long-memory" process with the focus of attention restricted on the "big" jumps in c. To do so, we would assume c_t to be a stochastic process with ρ-Hölder paths for $\rho \in (0, 1]$ where, up to some localization, we have $\mathbb{E}(|c_{S+s} - c_S| \mid \mathcal{F}_S) \leq K s^\rho$ for some constant K and all finite stopping times S. Doing so however would come at the cost of a considerable increase in the technical details of the paper.

3.1 Estimation and Asymptotic Theory the Spot Volatilities

The paper follows the now standard local averaging method to estimate the spot volatility. Let X and c be defined as in Section 2. Let k_n be a sequence of integers satisfying

$$k_n \to \infty \text{ and } k_n\Delta_n \to 0, \tag{6}$$

and let $u_n \asymp \Delta_n^{\varpi}$ be a sequence of truncation thresholds with $0 < \varpi < 1/2$. Define

$$\hat{c}(k_n, u_n)_i = \frac{1}{k_n\Delta_n} \sum_{m=0}^{k_n-1} \Delta_{i+m}^n X \Delta_{i+m}^n X^\top \mathbb{1}_{\{||\Delta_{i+m}^n X|| < u_n\}}. \tag{7}$$

Let $c_{t-} = \lim_{s\uparrow t} c_s$ then the left and right spot volatility estimators for c_t and c_{t-} for $t = i\Delta_n$ are defined as

$$\begin{aligned} \hat{c}_t &= \hat{c}(k_n, u_n)_{i+1} \\ \hat{c}_{t-} &= \hat{c}(k_n, u_n)_{i-k_n}. \end{aligned} \tag{8}$$

Notice that both the left and the right spot volatility estimators exclude the return corresponding to $t = i\Delta_n$.

The paper uses the following asymptotic theory for the spot volatility estimators which is due to Aït-Sahalia and Jacod (2014), Theorem 8.8. Let S_q be a sequence of finite stopping times. Under the models in (3) and (5), Assumptions 1 and 2, and if $k_n\sqrt{\Delta_n} \to \beta \in [0, \infty)$, we have the following finite-dimensional stable convergence in law

$$\begin{aligned} &\left(\sqrt{k_n}\left(\hat{c}(S_{q-}, k_n, u_n) - c_{S_{q-}}\right), \sqrt{k_n}\left(\hat{c}(S_q, k_n, u_n) - c_{S_q}\right)\right)_{q\geq 1} \\ &\quad \xrightarrow{\mathcal{L}-s} \left(U_{q-} + \beta U'_{q-}, U_q + \beta U'_q\right)_{q\geq 1} \end{aligned} \tag{9}$$

where the variables U_{q-}, U_q, U'_{q-}, and U'_q are defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and, conditionally on $\mathcal{F}$, are all independent centered Gaussian with covariances

$$\begin{aligned} \widetilde{\mathbb{E}}[U_{q-}^{ij} U_{q-}^{kl}] &= c_{S_{q-}}^{ik} c_{S_{q-}}^{jl} + c_{S_{q-}}^{il} c_{S_{q-}}^{jk}, \\ \widetilde{\mathbb{E}}[U_q^{ij} U_q^{kl}] &= c_{S_q}^{ik} c_{S_q}^{jl} + c_{S_q}^{il} c_{S_q}^{jk}, \\ \widetilde{\mathbb{E}}[U'^{ij}_{q-} U'^{kl}_{q-}] &= c_{S_{q-}}^{(c),ij,kl}, \\ \widetilde{\mathbb{E}}[U'^{ij}_{q} U'^{kl}_{q}] &= c_{S_q}^{(c),ij,kl}. \end{aligned} \tag{10}$$

Notice that the central limit theorem for estimating c_t (or c_{t-}) involves two random variables U_q and U'_q (or U_{q-} and U'_{q-}). This is because the estimation of c_t or c_{t-} involves two "errors": a "statical error" corresponding the random variables U_q and due to estimating the spot volatility using observed returns, and a "target error" corresponding the random variables U'_q.

To think about this target error define $E_{n,t} = \hat{c}_t - c_t = S_{n,t} + D_{n,t}$ where $S_{n,t}$ and $D_{n,t}$ are the statistical and target errors respectively. The random variables U_q and U'_q in (9) corresponding to limiting distributions of $S_{n,t}$ and $D_{n,t}$ respectively. The target error arises from the fact that, holding Δ_n fixed, $\hat{c}_t$ turns out to be an estimator of

$$\frac{1}{k_n\Delta_n}\int_{(i+1)\Delta_n}^{(i+k_n)\Delta_n} c_s ds \tag{11}$$

rather than c_t. It can be shown (see Aït-Sahalia and Jacod (2014), Section 8.1) that $S_{n,t}$ is of order $\sqrt{k_n}$ and $D_{n,t}$ is of order $(k_n\Delta_n)^{-1/2}$ which would imply that the best possible rate of convergence we could achieve would be $n^{1/4}$ by setting $k_n \asymp \Delta_n^{-1/2}$. The problem however is that the asymptotic distribution of the target error, corresponding to U'_q in (9), involves the volatility-of-volatility which is almost impossible to estimate in practice (see Aït-Sahalia and Jacod (2014) Section 8.3 for details). For this reason, the paper assumes from this point onwards that k_n is chosen so that $k_n\sqrt{\Delta_n} \to 0$ in which case $\beta = 0$ in (9) and the limiting distribution depends only on U_q and not U'_q. This can be achieved by taking $k_n \asymp \Delta_n^{-\rho}$ where $\rho \in (0, 1/2)$. Such a condition amounts to "under-smoothing" the estimates of c and makes $n^{1/4}$ an upper bound on the best possible rate of convergence one can achieve.

3.2 Estimating the Jump Times in the Market Volatility

As volatilities are latent processes, an important issue in the present setting is identifying the location of the jumps in the market volatility. A potential solution, motivated by the results in Todorov and Tauchen (2011), is to identify the jump times from a volatility index such as the Chicago Board of Options Exchange's (CBOE) Volatility Index (VIX).

To see how this might work let $d = 1$ in (3) and assume that X_t and c_t, in addition to their jump-intensities, are smooth functions of a Markovian state vector $\mathbf{S}_t$ which itself has a jump-diffusion representation. Whether explicitly stated as such or not this is a standard assumption in nearly all asset pricing models.[4] Next define the quadratic variation of X_t as $[X, X]_t$. A volatility index for X, such as the VIX, is defined as

$$v_t \equiv \mathbb{E}^{\mathbb{Q}}\left([X, X]_{t+N} - [X, X]_t | \mathcal{F}_t\right) \tag{12}$$

where $N > 0$ is fixed and $\mathbb{Q}$ is the risk-neutral measure. The volatility index v_t then is the expected change in the quadratic variation of X over the period $[t, t + N)$ under the risk-neutral measure. Under the setup in (3) and (5), Assumptions 1 and 2, and the above assumptions on a Markovian

[4] In fact, many models assume the spot-volatility itself is the sole state variable.

state vector $\mathbf{S}_t$ we can write v_t as a function of $\mathbf{S}_t$, i.e., $v_t = V(\mathbf{S}_t)$. Todorov and Tauchen (2011) show that under these assumptions that $V(\mathbf{S}_t)$ is a smooth function of $\mathbf{S}_t$. By Itô's formula for jump diffusions this implies that there is a jump in v_t if and only if there is a jump in $\mathbf{S}_t$ and, by extension, a jump in c_t. This idea is formalized in Proposition 1(b) below.

Given a volatility index v_t, such as the VIX, one can estimate the jumps in v_t using the now standard truncation based method introduced by Mancini (2001) as follows. To do so assume v_t has an Itô semi-martingale representation analogous to the ones for X_t and c_t in Assumptions 1 and 2 (formalized in Assumption 3(a) below). Let $(u_n)_{n\geq 1} \subset \mathbb{R}_+$ be a sequence of thresholds of the form

$$u_n \asymp \Delta_n^{\varpi} \text{ for some } \varpi \in (0, 1/2) \tag{13}$$

and define the set

$$\mathscr{I}_n^v = \{i : |\Delta_i^n v_t| > u_n\}. \tag{14}$$

Following Li, Todorov, and Tauchen (2015) it can be shown that $\mathscr{I}_n^v$ consistently locates the sampling intervals that contain jumps. That is,

$$\mathbb{P}(\mathscr{I}_n^v = \mathscr{I}_n^{v*}) \to 1 \tag{15}$$

where $\mathscr{I}_n^{v*} = \{i : \tau \in ((i-1)\Delta_n, i\Delta_n) \text{ for some } \tau \in \mathcal{T}^v\}$ and $\mathcal{T}^v = (\tau_p^v)_{p\geq 1}$ where $(\tau_p^v)_{p\geq 1}$ are the successive jump times of v_t.

The following proposition formalizes these arguments. Before presenting it however we need some additional notation. Let $c_{ZZ,t}$ denote the spot volatility of the market and define $\mathscr{I}_n = \{i : |\Delta_i^n c_{ZZ,t}| > u_n\}$ and $\mathscr{I}_n^* = \{i : \tau \in ((i-1)\Delta_n, i\Delta_n) \text{ for some } \tau \in \mathcal{T}\}$ where, as for v_t, we define $\mathcal{T} = (\tau_p)_{p\geq 1}$ where $(\tau_p)_{p\geq 1}$ are the successive jump times of $c_{ZZ,t}$.

Assumption 3. *Let v_t be defined as in* (12)*. We have the following assumptions.*

(a) Let

$$v_t = v_0 + \int_0^t b_s^{(v)} ds + \int_0^t \sigma_s^{(v)} dW_s^{(v)} + \delta^{(v)} * (\underline{p}^{(v)} - \underline{q}^{(v)})_t \tag{16}$$

where $b^{(v)}$ is locally bounded, $\sigma_t^{(v)}$ is bounded and nonzero for $t \in [0, T]$, $W^{(v)}$ is a 1-dimensional Brownian motion, $\underline{p}^{(v)}$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\underline{q}^{(v)}(dt, dx) = dt \otimes \lambda^{(v)}(dx)$, and $\delta^{(v)}$ is a predictable function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$. Finally, assume $\underline{q}^{(v)}([0, T], \mathbb{R}) < \infty$.

(b) Assume the following:

(i) $\boldsymbol{S}_t$ *is a* k*-dimensional vector with independent elements each of which solves, under* $\mathbb{Q}$*, the following*

$$df_t^{(i)} = \sum_{j=1}^{d_i} g_j^{(i)}\left(f_{t-}^{(i)}\right) dZ_{tj}^{(i)} \tag{17}$$

for $j = 1, \ldots, d_i$ *and* $i = 1, \ldots, k$ *where the functions* $g_j^{(i)}$ *are twice differentiable and* $Z_{tj}^{(i)}$ *are independent Lévy processes;*

(ii) $c_{ZZ,t} = G^{(c)}(\boldsymbol{S}_t)$ *for some twice differentiable function* $G^{(c)} : \mathbb{R}^k \to \mathbb{R}_+$ *with non-vanishing first derivatives on the support of* $\mathbb{S}_t$*;*

(iii) the compensator of the jumps in X_t*, i.e.,* $\underline{q}(dt, dx) = dt \otimes \lambda(dx)$ *in* (3)*, takes the form* $\underline{q}^{\mathbb{Q}}(dt, dx) = G^{(d)}(\boldsymbol{S}_t)dt \otimes \lambda^{\mathbb{Q}}(dx)$ *under* $\mathbb{Q}$ *where* $G^{(d)} : \mathbb{R}^k \to \mathbb{R}_+$ *is twice differentiable and* $\lambda^{\mathbb{Q}}$ *is a measure on* $\mathbb{R}$ *satisfying* $\int_{\mathbb{R}}(|x|^2 \wedge 1)\lambda^{\mathbb{Q}}(dx) < \infty$*.*

Part (a) of Assumption 3 establishes an analogous representation for v_t as is used for c_t. Nearly any jump-diffusion representation with a finite activity jump process would suffice. However, to maintain consistency we assume the representation in (16). Part (b)(i) introduces and formalizes the Markovian state vector $\mathbf{S}_t$ whereas parts (b)(ii)-(iii) introduce regularity conditions that will ensure v_t is continuously differentiable in $\mathbf{S}_t$.

Proposition 1. *Let* v_t *be defined as in* (12)*. We have the following results. (a) Under Assumption 3(a) we have* $\mathbb{P}(\mathscr{I}_n^v = \mathscr{I}_n^{v*}) \to 1$*. (b) Under Assumption 3(b)* $v_t = V(\boldsymbol{S}_t)$ *for some continuously differentiable function* $V : \mathbb{R}^k \to \mathbb{R}$*. Furthermore, if* $\partial V/\partial v_i \neq 0$ *on the support of* $\boldsymbol{S}_t$ *for* $i = 1, \ldots, k$ *and* V *is monotone in each of its arguments, then* $\mathcal{T}^v$ *and* $\mathcal{T}$ *coincide almost surely.*

Part (a) of Proposition 1 shows that the truncation based method presented here will correctly identify the jumps in the volatility index v_t. Part (b) shows, essentially, that if the spot volatility and the volatility index are continuously differentiable functions of a state vector then the jump times in the volatility index and the jump times in spot volatility of the market will coincide almost surely. Note that the proposition does not say that the volatility index v_t and the spot volatility of the market $c_{ZZ,t}$ are same rather it deals with the jump times in the two processes. In fact, because the volatility index is an expectation under the risk-neutral measure, the magnitude of jumps in the volatility index and in the spot volatility c_t need not, and will likely not, be the same. For this reason we will still need to estimate the jump sizes of the market volatility using the procedure in (7).

There is an additional complication that needs to be addressed in using the VIX as our volatility index. Andersen, Bondarenko, and Gonzalez-Perez (2015) document how the CBOE's method of

selecting the set of options to include in the calculation of the VIX leads to the inclusion of "spurious" jumps in the VIX. To guard against the inclusion of these "spurious" jumps in the empirical results, at the estimated jump times in the VIX, the corresponding changes in the spot volatility of the market are examined using an additional truncation based jump detection method to see if a jump in the spot volatility of the market was likely to have occurred or not.

4 Volatility Jump Regressions

This section details the estimation and theory behind jump regressions between two spot volatilities. The first subsection provides a framework for thinking about volatility jump regressions, a test for a constant volatility jump beta is introduced in the second subsection, the third subsection introduces a quasi-maximum-likelihood based estimator, and the final subsection introduces a bootstrap procedure for conducting inference.

4.1 A Framework for the Volatility Jump Beta

Let $X = (Z, Y)^\top$ and assume X and c are defined as in Section 2 and $\hat{c}_t$ and $\hat{c}_{t-}$ are estimated as in Section 3. When $X_t = (Z_t, Y_t)^\top$ we see that c_t takes the form

$$c_t = \begin{pmatrix} c_{ZZ,t} & c_{ZY,t} \\ c_{YZ,t} & c_{YY,t} \end{pmatrix}. \tag{18}$$

In many applications it proves useful to consider transformations of the spot volatility. To facilitate this let $g, h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be two continuous functions given by the researcher. To recover the spot volatilities of Z and Y we would set $g(c) = c_{YY}$ and $h(c) = c_{ZZ}$. An analysis of the spot volatilities in standard deviation or log units could be performed by setting $g(c) = \sqrt{c_{YY}}$ and $h(c) = \sqrt{c_{ZZ}}$ or $g(c) = \log c_{YY}$ and $h(c) = \log c_{ZZ}$ respectively. We might also consider $g(c) = c_{YY} - (c_{ZY})^2/c_{ZZ}$ which will recover the idiosyncratic volatility in Y. Let $(\tau_p)_{p\geq 1}$ be a set of jump times in $c_{ZZ,t}$. By the delta method we can extend the result in (9) to the current setting. That is,

$$\begin{aligned} &\sqrt{k_n}\left(g(\hat{c}_{\tau_p}) - g(c_{\tau_p}), g(\hat{c}_{\tau_{p-}}) - g(c_{\tau_{p-}}), h(\hat{c}_{\tau_p}) - h(c_{\tau_p}), h(\hat{c}_{\tau_{p-}}) - h(c_{\tau_{p-}})\right)_{p\geq 1} \\ &\quad \xrightarrow{\mathcal{L}-s} \left(\nabla g(c_{\tau_p})^\top U_{\tau_p}, \nabla g(c_{\tau_{p-}})^\top U_{\tau_{p-}}, \nabla h(c_{\tau_p})^\top U_{\tau_p}, \nabla h(c_{\tau_{p-}})^\top U_{\tau_{p-}}\right)_{p\geq 1} \end{aligned} \tag{19}$$

The (generalized) volatility jump regression considers the following model

$$\Delta g(c_{\tau_p}) = \beta_0(g, h)\Delta h(c_{\tau_p}) \tag{20}$$

for all $(\tau_p)_{p\geq 1}$ where $\Delta g(c_{\tau_p}) = g(c_{\tau_p}) - g(c_{\tau_p-})$ and $\Delta h(c_{\tau_p}) = h(c_{\tau_p}) - h(c_{\tau_p-})$. We can estimate the model in (20) as follows. Let $(i_p)_{p\geq 1}$ be the set of estimated jump times in the market volatility

following Section 3.2. The convergence of $(i_p)_{p\geq 1}$ to $(\tau_p)_{p\geq 1}$ follows from Proposition 1. Note that the rate convergence is "infinitely fast." Because of this the use of the estimated jump times $(i_p)_{p\geq 1}$ in (19) does not affect the asymptotic distributions. This implies

$$\begin{aligned}
&\sqrt{k_n}\left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p}), \Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p})\right)_{p\geq 1} \\
&\quad \xrightarrow{\mathcal{L}-s} \left(\nabla g(c_{\tau_p})^\top U_{\tau_p} + \nabla g(c_{\tau_{p-}})^\top U_{\tau_{p-}}, \nabla h(c_{\tau_p})^\top U_{\tau_p} + \nabla h(c_{\tau_{p-}})^\top U_{\tau_{p-}}\right)_{p\geq 1} \\
&\quad \equiv \left(\xi_{\tau_p}(g), \xi_{\tau_p}(h)\right).
\end{aligned} \tag{21}$$

Using these result we can write

$$\Delta g(\hat{c}_{i_p}) = \beta_0(g,h)\Delta h(\hat{c}_{i_p}) + \epsilon_{i_p}(g,h) \tag{22}$$

where

$$\epsilon_{i_p}(g,h) = \left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p})\right) - \beta_0(g,h)\left(\Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p})\right). \tag{23}$$

Together with (21) we see

$$\begin{aligned}
\sqrt{k_n}\epsilon_{i_p}(g,h) &\xrightarrow{\mathcal{L}-s} \xi_{\tau_p}(g) - \beta_0(g,h)\xi_{\tau_p}(h) \\
&\equiv \varphi_{\tau_p}(g,h)
\end{aligned} \tag{24}$$

for all $p \geq 1$. The results above provide the bulk of the intuition for the theorems to follow.

4.2 A Constant Volatility Beta Test

This subsection considers a test of the model in (20). That is a test of whether or not the constant volatility jump beta relationship in (20) holds at all jump times $(\tau_p)_{p\geq 1}$.[5] Formally, the testing problem is to decide which of the two sets the observed sample path falls within.

$$\begin{aligned}
&\Omega_0 \equiv \{\omega \in \Omega : \text{ condition (20) holds for some } \beta_0(g,h)(\omega) \text{ on path } \omega\}, \text{ or} \\
&\Omega_a \equiv \{\omega \in \Omega : \text{ condition (20) does not hold on path } \omega\}.
\end{aligned} \tag{25}$$

To do so consider the test statistic

$$Q_n(g,h) = \begin{pmatrix} Q_{ZZ,n} & Q_{ZY,n} \\ Q_{YZ,n} & Q_{YY,n} \end{pmatrix} = \sum_{p\geq 1} (\Delta g(\hat{c}_{i_p}), \Delta h(\hat{c}_{i_p}))^\top (\Delta g(\hat{c}_{i_p}), \Delta h(\hat{c}_{i_p})). \tag{26}$$

The result in (21) implies

$$\begin{aligned}
Q_n(g,h) &\xrightarrow{\mathbb{P}} \sum_{p\geq 1} (\Delta g(c_{\tau_p}), \Delta h(c_{\tau_p}))^\top (\Delta g(c_{\tau_p}), \Delta h(c_{\tau_p})) \\
&= \sum_{p\geq 1} \begin{pmatrix} \Delta g(c_{\tau_p})^2 & \Delta g(c_{\tau_p})\Delta h(c_{\tau_p}) \\ \Delta h(c_{\tau_p})\Delta g(c_{\tau_p}) & \Delta h(c_{\tau_p})^2 \end{pmatrix} \\
&\equiv Q(g,h).
\end{aligned} \tag{27}$$

[5] Out of necessity this paper only consider the case where $|(\tau_p)_{p\geq 1}| \geq 2$, i.e., $c_{ZZ,t}$ has at least two jumps over the interval $[0,T]$. This assumption is assumed to hold in all the results that follow.

By the Cauchy-Schwartz inequality notice that the condition in (20) is equivalent to the singularity of the matrix $Q(g,h)$. This implies a test for Ω_0, i.e., a test for a constant volatility jump beta, could be carried out by a one-sided test for whether $\det[Q(g,h)] = 0$. The determinant of $Q(g,h)$ can estimated by $\det[Q_n(g,h)]$. At a significance level $\alpha \in (0,1)$ the test will reject the null of a constant volatility jump beta if $\det[Q(g,h)] > cv_n^\alpha$ for a sequence of critical values cv_n^α described below.

Before specifying the critical values cv_n^α it is necessary to first discuss the asymptotic behavior of $\det[Q_n(g,h)]$ following the results in Section 4.1 above. Some algebra shows that in restriction to Ω_0 that

$$\det[Q_n(g,h)] = \left(\sum_{p\geq 1} \Delta h(\hat{c}_{i_p})^2\right)\left(\sum_{p\geq 1} \epsilon_{i_p}^2(g,h)\right) - \left(\sum_{p\geq 1} \Delta h(\hat{c}_{i_p})\epsilon_{i_p}(g,h)\right)^2. \tag{28}$$

This, together with the result in (21), implies

$$k_n \det[Q_n(g,h)] \xrightarrow{\mathcal{L}-s} \left(\sum_{p\geq 1} \Delta h(c_{\tau_p})^2\right)\left(\sum_{p\geq 1} \varphi_{\tau_p}^2(g,h)\right) - \left(\sum_{p\geq 1} \Delta h(c_{\tau_p})\varphi_{\tau_p}(g,h)\right)^2. \tag{29}$$

Using this result critical values cv_n^α can be constructed as follows.

Algorithm 1. *1. Simulate a collection of variables $(\tilde{U}_{i_p}, \tilde{U}_{i_{p-}})_{p\geq 1}$ consisting of copies of $(U_{\tau_p}, U_{\tau_{p-}})_{p\geq 1}$.*

2. Set

$$\begin{aligned} \tilde{\xi}_{i_p}(g) &= \nabla g(\hat{c}_{i_p})^\top \tilde{U}_{i_p} + \nabla g(\hat{c}_{i_{p-}})^\top \tilde{U}_{i_{p-}}, \\ \tilde{\xi}_{i_p}(h) &= \nabla h(\hat{c}_{i_p})^\top \tilde{U}_{i_p} + \nabla h(\hat{c}_{i_{p-}})^\top \tilde{U}_{i_{p-}}, \text{ and} \\ \tilde{\varphi}_{i_p}(g,h) &= \tilde{\xi}_{i_p}(g) - \bar{\beta}_n(g,h)\tilde{\xi}_{i_p}(h) \end{aligned} \tag{30}$$

where $\bar{\beta}_n(g,h) = Q_{ZY,n}(g,h)/Q_{ZZ,n}(g,h)$.

3. Compute

$$\tilde{\zeta}_n(g,h) = \left(\sum_{p\geq 1} \Delta h(\hat{c}_{i_p})^2\right)\left(\sum_{p\geq 1} \tilde{\varphi}_{i_p}^2(g,h)\right) - \left(\sum_{p\geq 1} \Delta h(\hat{c}_{i_p})\tilde{\varphi}_{i_p}(g,h)\right)^2. \tag{31}$$

4. Generate a large number of Monte Carlo simulations according to steps 1-3 and set cv_n^α as the $(1-\alpha)$-quantile of $\tilde{\zeta}_n(g,h)$ in the Monte Carlo sample.

Theorem 1 below provides the asymptotic justification for the proposed test. To state it define the following additional notation: recall $\xi_{\tau_p}(g)$ and $\xi_{\tau_p}(h)$ from (21) and set

$$\bar{\varphi}_{\tau_p}(g,h) = \xi_{\tau_p}(g) - \bar{\beta}(g,h)\xi_{\tau_p}(h) \tag{32}$$

for all $p \geq 1$ where $\bar{\beta}(g,h) = Q_{ZY}(g,h)/Q_{ZZ}(g,h)$. Note that in restriction to Ω_0 that $\bar{\varphi}_{\tau_p}(g,h) = \varphi_{\tau_p}(g,h)$.

Theorem 1. *Under the setup in* (3)-(5) *and Assumptions 1-2 and provided* $|(\tau_p)_{p\geq 1}| \geq 2$ *the following statements hold.*

(a). In restriction to Ω_0*, we have*

$$k_n \det[Q_n(g,h)] \xrightarrow{\mathcal{L}-s} \left(\sum_{p\geq 1} \Delta h(c_{\tau_p})^2\right)\left(\sum_{p\geq 1} \bar{\varphi}^2_{\tau_p}(g,h)\right) - \left(\sum_{p\geq 1} \Delta h(c_{\tau_p})\bar{\varphi}_{\tau_p}(g,h)\right)^2 \equiv \zeta(g,h). \tag{33}$$

(b). The sequence cv^a_n *of variables defined in Algorithm 1 converges in probability to the* $\mathcal{F}$*-conditional* $(1-\alpha)$*-quantile of* $\zeta(g,h)$*.*

(c). The test defined by the critical region $\{k_n \det[Q_n(g,h)] > cv^\alpha_n\}$ *has asymptotic size* α *under the null and asymptotic power one under the alternative. That is,*

$$\mathbb{P}\left(k_n \det[Q_n(g,h)] > cv^\alpha_n | \Omega_0\right) \to \alpha \quad \textit{and} \quad \mathbb{P}\left(k_n \det[Q_n(g,h)] > cv^\alpha_n | \Omega_a\right) \to 1. \tag{34}$$

Part (a) of Theorem 1 describes the stable convergence of the test statistic $\det[Q_n(g,h)]$ under the null hypothesis, which occurs at rate k_n. The limiting variable $\zeta(g,h)$ is quadratic in the variables $\bar{\varphi}_{\tau_p}(g,h)$ which, conditional on $\mathcal{F}$, are mutually independent mixed Gaussian variables with a covariance structure derivable from (10). Comparing (31) and (33), it is easy to see $\tilde{\zeta}_n(g,h)$ is designed to mimic the limiting variable $\zeta(g,h)$. Part (b) shows the quantile of $\tilde{\zeta}_n(g,h)$ will consistently estimate that of $\zeta(g,h)$. Note that part (b) holds under both the null and the alternative. Part (c) shows that the proposed test has valid size control and is consistent against general fixed alternatives.

4.3 Estimating the Volatility Jump Beta

This subsection consider the estimation of the volatility jump beta, i.e., the estimation of $\beta_0(g,h)$ in (20). The ordinary least squares (OLS) estimator, while asymptotically valid, will turn out to have poor small sample properties due to the estimation error in estimating both $\Delta g(\hat{c})$ and $\Delta h(\hat{c})$. To correct for this problem this paper proposes a quasi-maximum-likelihood approach based on the asymptotic distribution in (24).

To proceed define the OLS estimator of $\beta_0(g,h)$ as

$$\hat{\beta}^{OLS}_n(g,h) = \frac{\sum_{p\geq 1} \Delta h(\hat{c}_{i_p})\Delta g(\hat{c}_{i_p})}{\sum_{p\geq 1}(\Delta h(\hat{c}_{i_p}))^2}. \tag{35}$$

By (21) we see that asymptotically

$$\hat{\beta}_n^{OLS}(g,h) \xrightarrow{\mathbb{P}} \frac{\sum_{p\geq 1}\Delta h(c_{\tau_p})\Delta g(c_{\tau_p})}{\sum_{p\geq 1}(\Delta h(c_{\tau_p}))^2} = \beta_0(g,h) \tag{36}$$

as $k_n \to \infty$ and $k_n\sqrt{\Delta_n} \to 0$. However, while this remains an asymptotically valid approach to estimating $\beta_0(g,h)$, in small samples the estimator will likely be biased due to the estimation error in estimating both $\Delta g(\hat{c})$ and $\Delta h(\hat{c})$. The intuition for this result follows analogously to problems of "attenuation bias" or the classical "errors-in-variables" problem. These problems are present in any ordinary least squares estimation, but are especially acute here due to the high estimation error in estimating the spot volatilities.

To see this consider the following heuristic model. Let

$$\Delta g(c_{\tau_p}) = \beta_0(g,h)\Delta h(c_{\tau_p}) \tag{37}$$

as before and assume

$$\begin{aligned} \Delta g(\hat{c}_{i_p}) &= \Delta g(c_{\tau_p}) + e_{i_p}(g), \\ \Delta h(\hat{c}_{i_p}) &= \Delta h(c_{\tau_p}) + e_{i_p}(h) \end{aligned} \tag{38}$$

for all $p \geq 1$ where $e_{i_p}(g)$ and $e_{i_p}(g)$ are independent with $e_{i_p}(g) \overset{iid}{\sim} N(0,\sigma^2_{g,i_p})$ and $e_{i_p}(h) \overset{iid}{\sim} N(0,\sigma^2_{h,i_p})$. In this heuristic setting we have

$$\hat{\beta}_n^{OLS}(g,h) = \frac{\widehat{Cov}(\Delta g(\hat{c}_{i_p}),\Delta h(\hat{c}_{i_p}))}{\widehat{Var}(\Delta h(\hat{c}_{i_p}))} \tag{39}$$

and in expectation

$$\begin{aligned} Cov(\Delta g(\hat{c}_{i_p}),\Delta h(\hat{c}_{i_p})) &= \beta_0(g,h)Var(\Delta h(c_{\tau_p})) \\ Var(\Delta h(\hat{c}_{i_p})) &= Var(\Delta h(c_{\tau_p})) + \sigma^2_{h,i_p}. \end{aligned} \tag{40}$$

Which implies that $\hat{\beta}_n^{OLS}(g,h)$ is actually an estimator of

$$\beta_0(g,h)\frac{Var(\Delta h(c_{\tau_p}))}{Var(\Delta h(c_{\tau_p})) + \sigma^2_{h,i_p}} \tag{41}$$

rather than $\beta_0(g,h)$.

To motivate an alternative estimator recall the results in (19)-(24) where it was shown that

$$\Delta g(\hat{c}_{i_p}) = \beta_0(g,h)\Delta h(\hat{c}_{i_p}) + \epsilon_{i_p}(g,h) \tag{42}$$

and $\sqrt{k_n}(\epsilon_{i_p}(g,h))_{p\geq 1} \xrightarrow{\mathcal{L}-s} (\varphi_{\tau_p}(g,h))_{p\geq 1}$ where $\varphi_{\tau_p}(g,h))_{p\geq 1}$ are independent Gaussian. A potential estimator could be based on this result, i.e. the result that $\epsilon_{i_p}(g,h)$ are asymptotically Gaussian and independent. The reason for such a proposal is that a likelihood based approach can correct

for the attenuation bias or the errors-in-variables problem by considering the estimating error in both the independent and dependent variables of the regression. Ordinary least squares in contrast is limited because, by solely minimizing the residuals of the dependent variable, it cannot take into account the estimation error in the independent variable.

What this paper proposes then is a quasi-maximum-likelihood estimator of $\beta(g,h)$ based on an asymptotic approximation of the limiting distribution of $\epsilon_{i_p}(g,h)$. To do so, recall $\varphi_{\tau_p}(g,h) = \xi_{\tau_p}(g) - \beta_0(g,h)\xi_{\tau_p}(h)$ where $(\xi_{\tau_p}(g), \xi_{\tau_p}(h))$ were defined in (21). Denote the covariances of $(\xi_{\tau_p}(g), \xi_{\tau_p}(h))$ as Σ_{τ_p} and note that by (10) we can express $\Sigma_{\tau_p} = \Sigma(g,h,c_{\tau_{p-}}, c_{\tau_{p+}})$. Closed form solutions for Σ_{τ_p} are easy to derive from (10) given functional forms of g and h. With Σ_{τ_p} so defined the variances of $\varphi_{\tau_p}(g,h)$ take the form $\Sigma^{11}_{\tau_p} + \beta_0(g,h)^2\Sigma^{22}_{\tau_p} - 2\beta_0(g,h)\Sigma^{12}_{\tau_p}$ for all $p \geq 1$. This gives us a limiting log-likelihood for (42) of the form

$$\log \mathcal{L}\left(b|\Sigma_{\tau_p}\right) = -\frac{1}{2}\sum_{p\geq 1}\log\left(2\pi(\Sigma^{11}_{\tau_p} + b^2\Sigma^{22}_{\tau_p} - 2b\Sigma^{12}_{\tau_p})\right) - \sum_{p\geq 1}\frac{\left(\Delta g(c_{\tau_p}) - b\Delta h(c_{\tau_p})\right)^2}{2\left(\Sigma^{11}_{\tau_p} + b^2\Sigma^{22}_{\tau_p} - 2b\Sigma^{12}_{\tau_p}\right)}. \tag{43}$$

However, an estimator based on the likelihood in (43) is infeasible since the likelihood is based on the limiting variables $\varphi_{\tau_p}(g,h)$ and Σ_{τ_p}. Instead we can use an approximating log-likelihood of the form

$$\begin{aligned}\log \hat{\mathcal{L}}_n\left(b|\hat{\Sigma}_{i_p}\right) = &-\frac{1}{2}\sum_{p\geq 1}\log\left(2\pi k_n^{-1}(\hat{\Sigma}^{11}_{i_p} + b^2\hat{\Sigma}^{22}_{i_p} - 2b\hat{\Sigma}^{12}_{i_p})\right) \\ &- \sum_{p\geq 1}\frac{\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)^2}{2k_n^{-1}\left(\hat{\Sigma}^{11}_{i_p} + b^2\hat{\Sigma}^{22}_{i_p} - 2b\hat{\Sigma}^{12}_{i_p}\right)}.\end{aligned} \tag{44}$$

where $\hat{\Sigma}_{i_p} = \Sigma(g,h,\hat{c}_{i_{p-}}, \hat{c}_{i_{p+}})$. The feasible quasi-ML estimator then is the arg-max of the approximating log-likelihood above, i.e.,

$$\hat{\beta}^{MLE}_n(g,h) = \arg\max_{b\in\mathbb{R}} \log \hat{\mathcal{L}}_n\left(b|\hat{\Sigma}_{i_p}\right). \tag{45}$$

The following theorem establishes the asymptotic properties of $\hat{\beta}^{MLE}_n(g,h)$. To do define the following. Let

$$\begin{aligned}\Xi(g,h) \equiv &\frac{1}{2}\sum_{p\geq 1}\frac{\varphi_{\tau_p}(g,h)(\Delta g(c_{\tau_p}) - \beta_0\Delta h(c_{\tau_p}))\left(2\beta_0\Sigma^{22}_{\tau_p} - 2\Sigma^{12}_{\tau_p}\right)}{\left(\Sigma^{11}_{\tau_p} + \beta_0^2\Sigma^{22}_{\tau_p} - 2\beta_0\Sigma^{12}_{\tau_p}\right)^2} \\ &+ \sum_{p\geq 1}\frac{\Delta h(c_{\tau_p})\varphi_{\tau_p}(g,h)}{\Sigma^{11}_{\tau_p} + \beta_0^2\Sigma^{22}_{\tau_p} - 2\beta_0\Sigma^{12}_{\tau_p}}\end{aligned} \tag{46}$$

and

$$H(g,h) \equiv \sum_{p\geq 1} -\frac{2\Delta h(c_{\tau_p})(\Delta g(c_{\tau_p}) - \beta_0 \Delta h(c_{\tau_p}))(2\beta_0 \Sigma^{22}_{\tau_p} - 2\Sigma^{12}_{\tau_p})}{\left(\Sigma^{11}_{\tau_p} + \beta_0^2 \Sigma^{22}_{\tau_p} - 2\beta_0 \Sigma^{12}_{\tau_p}\right)^2}$$
$$- \frac{1}{2}(\Delta g(c_{\tau_p}) - \beta_0 \Delta h(c_{\tau_p}))^2 \left(\frac{2(2\beta_0 \Sigma^{22}_{\tau_p} - 2\Sigma^{12}_{\tau_p})^2}{\left(\Sigma^{11}_{\tau_p} + \beta_0^2 \Sigma^{22}_{\tau_p} - 2\beta_0 \Sigma^{12}_{\tau_p}\right)^3} - \frac{2\Sigma^{22}_{\tau_p}}{\left(\Sigma^{11}_{\tau_p} + \beta_0^2 \Sigma^{22}_{\tau_p} - 2\beta_0 \Sigma^{12}_{\tau_p}\right)^2} \right) \quad (47)$$
$$- \frac{\left(\Delta h(c_{\tau_p})\right)^2}{\Sigma^{11}_{\tau_p} + \beta_0^2 \Sigma^{22}_{\tau_p} - 2\beta_0 \Sigma^{12}_{\tau_p}}$$

where for the ease of notation let $\beta_0 = \beta_0(g,h)$ above.

Theorem 2. *Under Assumptions 1 and 2, provided g and h are both continuously differentiable, and provided $|(\tau_p)_{p\geq 1}| \geq 2$ the following statements hold. We have*

(a). $\hat{\beta}_n^{MLE}(g,h) \xrightarrow{\mathbb{P}} \beta_0(g,h)$

(b). $\sqrt{k_n}\left(\hat{\beta}_n^{MLE}(g,h) - \beta_0(g,h)\right) \xrightarrow{\mathcal{L}-s} -(H(g,h))^{-1}\Xi(g,h)$

as $k_n \to \infty$ and $k_n\sqrt{\Delta_n} \to 0$.

Note that the theorem holds under both Ω_0 and Ω_a. In restriction to Ω_a one could think of $\beta_0(g,h)$ as the "pseudo-true parameter" being the arg-max of the likelihood in (43) without its implications on the volatility jump regression framework in (20). Notice that not only is the distribution in Theorem 2(b) highly non-standard but that it depends on the latent variables β_0, c_{τ_p-}, and c_{τ_p+}. For these reasons this paper proposes a bootstrap method for conducting inference below in Section 4.4.

4.4 Bootstrap Based Inference

As the distribution of $\hat{\beta}_n^{MLE}(g,h)$ is highly non-standard this paper follows Li, Todorov, Tauchen, and Chen (2016) and introduces a bootstrap procedure to provide feasible inference followed by a justification of its asymptotic validity. The bootstrap procedure was first introduced to the high-frequency setting by Goncalves and Meddahi (2009) for making inference about the integrated variance and covariance matrices and is conceptually simple to grasp. The researcher need only repeatedly compute estimators in the bootstrap samples following the algorithm below. Other papers that have used the bootstrap in a high-frequency setting are Dovonon, Gonçalves, and Meddahi (2013), Dovonon, Gonçalves, Hounyo, and Meddahi (2014), and Hounyo (2013). Note that this paper uses a form of the i.i.d. bootstrap in Goncalves and Meddahi (2009) as opposed to

a bootstrap based on the local Gaussianity of returns as in Li, Todorov, Tauchen, and Chen (2016) and introduced in Hounyo (2013).

Algorithm 2. *Let k_n be the number of returns in the local averaging window of the spot volatility estimator as before.*

1. *For each $\{i_p\}_{p\geq 1}$ re-sample, with replacement, the k_n returns before and after each jump time. That is, re-sample from the sets*

$$\begin{aligned} &\{\Delta_j^n X\}_j \quad \text{for } j = i_p - k_n, \ldots, i_p - 1, \text{ and} \\ &\{\Delta_k^n X\}_k \quad \text{for } k = i_p + 1, \ldots, i_p + k_n. \end{aligned} \tag{48}$$

   Denote the resampled sets as $\{\Delta_j^n X^*\}_j$ and $\{\Delta_k^n X^*\}_k$ where $j = i_p - k_n, \ldots, i_p - 1$ and $k = i_p + 1, \ldots, i_p + k_n$ for each $\{i_p\}_{p\geq 1}$.

2. *For each bootstrap sample $\{\Delta_i^n X^*\}$ compute $\Delta g(\hat{c}^*_{i_p})$ and $\Delta h(\hat{c}^*_{i_p})$ as in (7).*

3. *Compute $\hat{\beta}_n^{*MLE}(g,h)$ using $\Delta g(\hat{c}^*_{i_p})$ and $\Delta h(\hat{c}^*_{i_p})$ for all $\{i_p\}_{p\geq 1}$.*

In summary, Algorithm 2 suggests computing $\hat{\beta}_n^{*MLE}(g,h)$ in the same manner as $\hat{\beta}_n^{MLE}(g,h)$ using the bootstrap sample. Theorem 3 below shows that the bootstrap spot volatility estimates $(\hat{c}^*_{i_{p-}}, \hat{c}^*_{i_{p+}})_{p\geq 1}$ converge in probability to $(c_{\tau_{p-}}, c_{\tau_{p+}})_{p\geq 1}$ and describes the convergence in probability of the $\mathcal{F}$-conditional law of the bootstrap estimator $\hat{\beta}_n^{*MLE}(g,h)$.

Theorem 3. *Suppose the same conditions as Theorem 2.*

(a). We have

$$\left(\hat{c}^*_{i_{p-}}, \hat{c}^*_{i_{p+}}\right)_{p\geq 1} = \left(c_{\tau_{p-}}, c_{\tau_{p+}}\right)_{p\geq 1} + o_p(1), \tag{49}$$

as $k_n \to \infty$ and $k_n\sqrt{\Delta_n} \to 0$.

(b). We can decompose

$$\sqrt{k_n}\left(\hat{\beta}_n^{*MLE}(g,h) - \hat{\beta}_n^{MLE}(g,h)\right) = -\left(H_n^*(g,h)\right)^{-1}\Xi_n^*(g,h) + o_p(1), \tag{50}$$

such that

$$(H_n^*(g,h), \Xi_n^*(g,h)) \xrightarrow{\mathcal{L}|\mathcal{F}} (H(g,h), \Xi(g,h) \tag{51}$$

where $(H(g,h), \Xi(g,h))$ are defined as in (46) and (47).

Theorem 3 justifies using the $\mathcal{F}$-conditional distribution $\sqrt{k_n}(\hat{\beta}_n^{*MLE}(g,h) - \hat{\beta}_n^{MLE}(g,h))$ to approximate the $\mathcal{F}$-conditional distribution $\sqrt{k_n}(\hat{\beta}_n^{MLE}(g,h) - \beta_0(g,h))$. Note that the limiting bootstrap distribution in Theorem 3 is not symmetric. Because of this two-sided confidence intervals computed using methods designed to place equal probability on either end of the confidence interval, such as the basic or simple bootstrap, will not in general result in confidence intervals symmetric around the estimated volatility jump beta.

5 Simulation Results

This section examines the performance of the proposed techniques on simulated data in a series of Monte Carlo experiments designed to mimic the empirical setting in Section 6. The study examines a volatility jump regression setting in log units, that is $g(c) = \log c_{YY}$ and $h(c) = \log c_{ZZ}$. The sample span is set as $T = 252$ to match the trading days in a typical year. The study considers both the case when $n = 78$ corresponding to five-minute sampling and the case when $n = 390$ corresponding the one-minute sampling. When $n = 78$ the local averaging windows are of size $k_n = 10$, 15, and 20 and when $n = 390$ the local averaging windows are of size $k_n = 20$, 30, 40 and 50. There are 1000 Monte Carlo trials.

The study uses the following data generating process

$$\begin{aligned} d\log c_{ZZ,t} &= \rho_Z(\mu_Z - \log c_{ZZ,t-1}) + \sigma_Z dB_{Z,t} + J_t dN_t \\ d\log c_{YY,t} &= \rho_Y(\mu_Y - \log c_{YY,t-1}) + \sigma_Y dB_{Y,t} + \beta_t J_t dN_t \end{aligned} \tag{52}$$

and

$$\begin{aligned} dZ_t &= \sqrt{c_{ZZ,t}} dW_{Z,t} \\ dY_t &= \sqrt{c_{YY,t}} dW_{Y,t} \end{aligned} \tag{53}$$

where $\{B_{Z,t}, B_{Y,t}, W_{Z,t}, W_{Y,t}\}$ are independent Brownian, N_t is a Poisson process with intensity λ, and

$$J_t \overset{iid}{\sim} N(0, \phi^2). \tag{54}$$

The data generating process necessarily includes volatility co-jumps. Price-volatility co-jumps could be included but would not materially impact the results. This is because the estimation of the jump in the spot-volatility only includes returns before and after the jump time in the volatility and not the return that overlaps the volatility jump.

The volatility jump beta is set as

$$\begin{cases} \beta_t = \beta_0 \text{ for all } t \in [0,T] & \text{under } H_0 \\ d\beta_t = \rho_\beta(1 - \beta_{t-1}) + \sigma_\beta \sqrt{\beta_{t-1}} dB_{\beta,t} & \text{under } H_a \end{cases} \tag{55}$$

Table 1: Monte Carlo Rejection Rates for the Constant Volatility Jump Beta Test

			Under H_0			Under H_a		
	n	k_n	1%	5%	10%	1%	5%	10%
$\sigma = 0.1$	78	10	1.4	4.7	8.6	44.5	54.1	60.3
	78	15	0.9	5.5	9.9	55.6	64.7	70.5
	78	20	1.0	6.1	11.5	63.1	71.5	77.0
	390	20	1.3	5.1	8.8	63.7	71.1	75.9
	390	30	1.3	5.0	8.9	72.4	79.6	83.2
	390	40	1.2	5.0	10.8	79.1	84.7	87.1
	390	50	1.8	5.4	10.2	82.8	87.2	90.0
$\sigma = 0.2$	78	10	1.0	4.4	8.4	45.8	56.4	61.6
	78	15	0.8	4.5	10.4	55.7	64.3	69.6
	78	20	1.4	5.6	12.2	64.4	70.8	75.9
	390	20	1.0	4.6	10.0	62.8	70.8	75.3
	390	30	0.9	5.5	10.4	72.4	79.7	83.9
	390	40	1.5	6.8	13.3	79.3	84.7	87.4
	390	50	1.6	6.2	12.8	82.8	87.9	90.2
$\sigma = 0.3$	78	10	1.4	5.2	9.3	43.4	54.2	59.9
	78	15	1.7	6.2	11.6	54.2	63.9	69.6
	78	20	2.8	6.8	12.4	62.3	71.0	75.0
	390	20	1.3	5.5	11.2	63.5	71.9	75.8
	390	30	1.5	6.7	12.7	73.3	81.2	84.7
	390	40	1.6	7.2	13.2	81.6	87.0	88.9
	390	50	2.9	9.4	15.6	84.9	89.0	91.0

Note: Reported are the Monte Carlo rejection rates of the constant volatility jump beta test at significance levels of 1%, 5%, and 10%. Results are given both under the null hypothesis (H_0) and under the alternative hypothesis (H_a) for various sampling schemes and local estimation windows ($k_n = 10$, 15, and 20 when $n = 78$ and $k_n = 20$, 30, 40, and 50 when $n = 390$) as well as various volatility-of-volatility specifications ($\sigma = 0.1$, 0.2, and 0.3). Each experiment had 1000 Monte Carlo trials.

Table 2: Monte Carlo Volatility Jump Beta Estimation Results and Confidence Interval Coverages

			OLS				MLE			
	n	kn	Bias	99%	95%	90%	Bias	99%	95%	90%
$\sigma = 0.1$	78	10	-0.306	73.6	50.0	34.6	-0.184	98.5	92.3	85.8
	78	15	-0.250	84.2	63.2	49.0	-0.116	98.5	93.9	88.2
	78	20	-0.224	89.3	70.5	56.2	-0.078	99.0	96.2	91.7
	390	20	-0.164	82.9	61.8	48.6	-0.132	96.7	88.9	80.9
	390	30	-0.123	88.7	71.5	59.3	-0.090	97.8	91.1	84.9
	390	40	-0.105	92.8	78.7	66.6	-0.070	98.5	94.3	88.0
	390	50	-0.094	93.9	81.5	71.7	-0.057	99.1	96.1	90.4
$\sigma = 0.2$	78	10	-0.293	74.3	49.4	34.6	-0.172	98.7	93.2	87.5
	78	15	-0.242	86.2	62.5	46.2	-0.116	99.1	95.2	90.6
	78	20	-0.218	90.0	71.0	56.3	-0.077	99.4	96.8	92.3
	390	20	-0.174	80.8	57.3	44.1	-0.133	96.7	87.8	80.1
	390	30	-0.132	87.4	70.0	58.1	-0.094	97.1	90.5	84.3
	390	40	-0.109	91.5	74.5	63.0	-0.069	98.1	92.6	86.1
	390	50	-0.095	92.9	78.2	67.0	-0.052	98.7	94.2	88.4
$\sigma = 0.3$	78	10	-0.289	79.0	54.2	41.0	-0.158	98.6	93.1	86.6
	78	15	-0.235	89.5	66.4	51.9	-0.103	98.7	95.4	90.5
	78	20	-0.205	91.4	75.6	61.1	-0.062	99.1	96.0	93.1
	390	20	-0.168	83.9	63.9	50.7	-0.132	96.4	89.4	81.8
	390	30	-0.129	90.0	73.9	60.3	-0.094	97.4	92.0	86.1
	390	40	-0.109	91.6	78.8	67.6	-0.069	98.4	93.1	87.6
	390	50	-0.098	93.8	81.9	70.6	-0.055	98.5	94.7	90.3

Note: Reported are the mean bias across Monte Carlo trials and the Monte Carlo coverage rates for the confidence intervals (CIs) at levels 99%, 95%, and 90%. Results are given for various sampling schemes and local estimation windows ($k_n = 10$, 15, and 20 when $n = 78$ and $k_n = 20$, 30, 40, and 50 when $n = 390$) as well as various volatility-of-volatility specifications ($\sigma = 0.1$, 0.2, and 0.3). The CIs are constructed using Algorithm 2 and the basic or simple bootstrap based on 1000 bootstrap draws. Each experiment had 1000 Monte Carlo trials.

where $B_{\beta,t}$ is a Brownian motion independent of $\{B_{Z,t}, B_{Y,t}, W_{Z,t}, W_{Y,t}\}$. Under this setup the unconditional mean of β_t under the alternative is the same as β_0.

The following parameter specifications are used

$$\begin{cases} \mu_Z = 0.05, \mu_Y = 0.55, & \text{with } \log c_{ZZ,0} = \mu_Z \text{ and } \log c_{YY,0} = \mu_Y, \\ \phi = 1.1 & \\ \lambda = 15/252 & \\ \sigma_Z = \sigma_Y = \sigma, & \text{with } \sigma\text{=0.1, 0.2, or 0.3,} \\ \rho_Z = \rho_Y = 0.15 & \\ \beta_0 = 1, & \text{with } \rho_\beta = 0.15 \text{ and } \sigma_\beta = 0.7. \end{cases} \tag{56}$$

The parameter σ controls the volatility-of-volatility and thereby the signal-to-noise ratio of the estimated volatility jump beta. This parameter is varied over three different specifications. While the empirical volatility-of-volatility is almost impossible to estimate in practice these specifications led to realistic paths for the spot volatilities. The jump intensity λ is set so that there should be on average 15 volatility jumps in each sample which is close to what is observed in the data. The parameters μ_Z and μ_Y determine the means of c_Z and c_Y. With $\mu_Z = 0.05$ and $\mu_Y = 0.55$ the mean annualized volatility of Z is approximately 16.2 and the mean annualized volatility of Y is approximately 20.8. The parameters ρ_Z and ρ_Y control the degree of mean reversion in the spot volatilities. With $\rho_Z = \rho_Y = 0.15$ the spot volatilities have a half-life of approximately 4.6 days. Finally, the parameter ϕ controls the size of the log-volatility jumps. Given the log-volatility setting with $\phi = 1.1$ the volatility jumps have a standard deviation of 1.1 percent.

Table 1 reports the finite-sample rejection rates for the constant volatility jump beta test as described in Theorem 1. In general, the test appears to have good size and power properties. Under the null-hypothesis the rejection rates are fairly close to their nominal levels though they increase both as k_n increases and σ increases. Under the alternative the rejection rates are well above their nominal levels. The increase in the rejection rates as k_n increases is due to the target error described in Section 3.1. Recall that for a fixed n the target error grows with k_n. While asymptotically dominated by the statistical error the target remains in finite-samples and necessitates the need to "under-smooth" our estimates, i.e. to choose a k_n smaller than might be suggested solely by the asymptotics for the statistical error. In particular using $k_n = 10$ or 15 when $n = 78$ and using $k_n = 30$ when $n = 390$ appears to mitigate the effect of the target error in this context while maintaining the asymptotic convergence properties of the test. The increase in the rejection rates under the null as σ increases is due to the increased volatility-of-volatility as σ increases which in turn exacerbates the target error.

Table 2 reports summary statistics for the ordinary least squares (OLS) estimator and the

quasi-maximum-likelihood estimator (MLE) of the volatility jump beta and coverage results for the confidence intervals. The OLS and ML estimators are detailed in Section 4.3 and the confidence intervals were constructed using the bootstrap procedure of Algorithm 2 detailed in Section 4.4 and the basic or simple bootstrap confidence interval procedure with 1000 replications. Other bootstrap confidence intervals, such as the bias corrected and acceleration adjusted (BC_a) confidence intervals in Efron (1987), were tried but were found to have worse coverage properties than the basic bootstrap confidence intervals. (For an overview of different bootstrap confidence interval methods see DiCiccio and Efron (1996) or Davison and Hinkley (1997).) In terms of the mean bias of the estimators in Table 2 the results quite strongly support the hypothesis of Section 4.3 where it was proposed that the OLS estimator would be significantly biased downwards by an attenuation or an errors-in-variables bias due the estimation error in estimating both the spot volatility jumps of the market and the asset. The biases for the ML estimator, while not completely eliminated, are significantly smaller in every case supporting its use over the OLS estimator. The coverage results for the ML estimator are significantly closer to their nominal levels compared to the OLS estimator which is again consistent again with the theory that the OLS estimator is significantly biased. In light of the poor coverage result and the biases this paper does not believe the OLS estimator should be used to conduct inference.

6 Empirical Applications

This section contains three empirical applications. The first on estimating and conducting inference on a series of volatility jump betas, the next on an examination of the jump beta for the idiosyncratic volatility, and the last on the pricing of volatility jumps. The data used in these studies and the specifics in how the algorithms and theorems of the paper are implemented are detailed in the subsection below.

6.1 Data and Implementation

The empirical applications in this section use two sets of data. The first set of data consist of the ETFs on the nine industry portfolios comprising the S&P500 index. These portfolios are, with ticker symbols in parenthesis, as follows: materials (XLB), energy (XLE), financials (XLF), industry (XLI), technology (XLK), consumer staples (XLP), utilities (XLU), healthcare (XLV), and consumer discretionary (XLY). The second set of data are the thirty stocks comprising the Dow Jones Industrial Average Index (DOW30) as of December 2015 with the exception that Visa (V) is replaced with Bank of America (BAC) to make a balanced panel covering January 3, 2007

to December 21, 2015. The proxy for the market is the front-month E-mini S&P500 index futures contract (ES), which is among the most liquid assets in the world. Market holidays and half trading days are removed. The two Flash Crashes (May 6, 2010 and April 23, 2013) are removed as well because the dramatic market fluctuations in these days are known to be due to market malfunctioning. The resultant sample contains 2225 trading days. The intraday observations are sampled at a one-minute frequency from 9:35 to 15:55 EST; the prices at the first and the last five minutes are discarded to guard against possible adverse microstructure effects near the market opening and closing and to avoid issues relating to the opening and closing auctions. In the predictive regressions of Section 6.4 the three-month treasury bill rate is used to proxy for the risk free rate. As in the Monte Carlo section the analysis is done in log units and, due to the poor coverage and bias results of the OLS estimator in the Monte Carlo study of Section 5, only results for the ML estimator are provided. In addition, the local averaging window is set to $k_n = 30$.

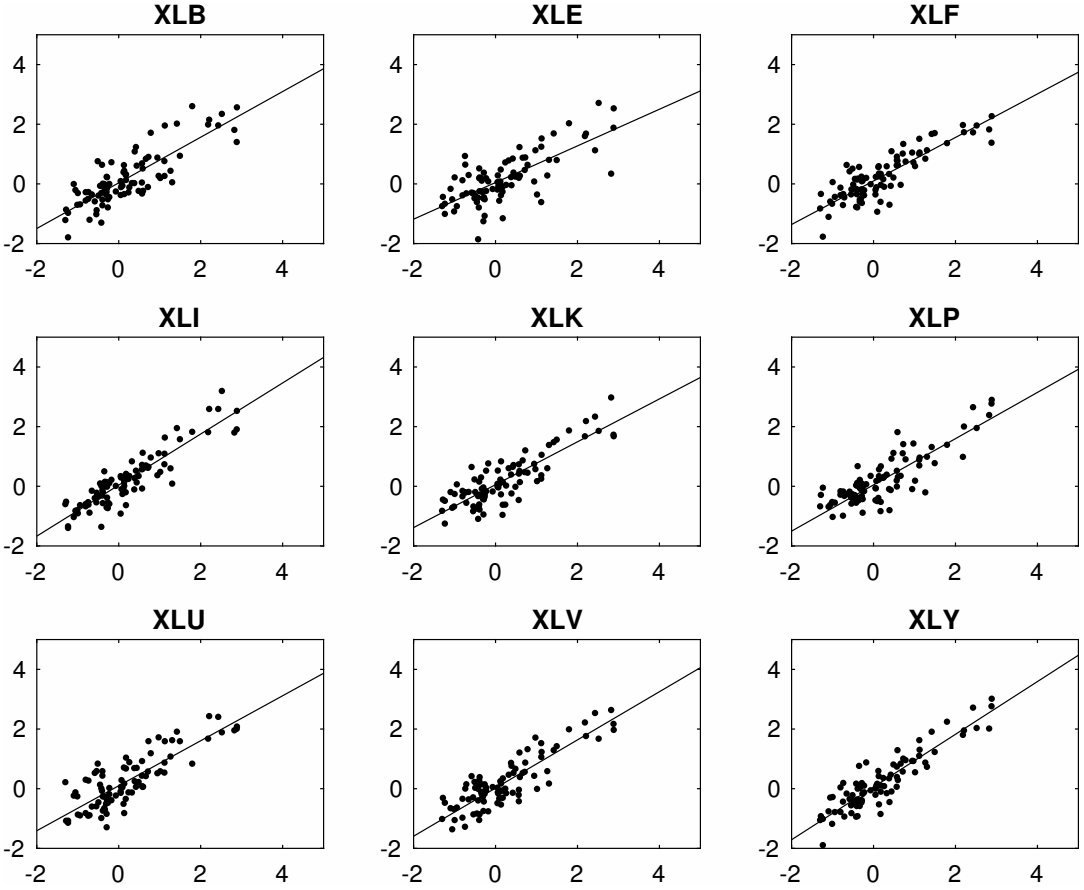
The estimation of the jumps in the market volatility follows Section 3.2 and uses the Chicago Board of Options Exchange's (CBOE) Volatility Index (VIX). To be especially careful to guard against the inclusion of "spurious" jumps, the VIX is sampled at a five-minute frequency and the jump truncation scheme used to identify jumps is set for moves slightly larger in magnitude than six local standard deviations. As in prior work, the truncation threshold is also scaled to account for the deterministic diurnal volatility pattern, but the details are omitted for brevity. (See the supplementary material of Todorov and Tauchen (2012) for details on the procedure.) These jump times were then verified against the corresponding changed in the spot volatility of the E-mini as explained in Section 3.2. Together these procedures led to the detection of 97 market volatility jumps across the whole sample.

6.2 Volatility Jump Regressions

This set of empirical applications estimates the volatility jump beta and performs tests of a constant volatility jump beta for both the S&P Sector ETFs and the DOW30 stocks described above. Figures 2 and 3 display scatter plots of the estimated spot volatility jumps in the S&P industry portfolios and the DOW30 stocks against the estimated spot volatility jumps in the E-mini. These scatter plots are estimates, in a sense, of the proposition that a constant volatility jump beta exists. The figures also show the line of best fit based on the ML estimator developed in Section 4.3. While the plots do not display a perfectly linear fit this not be expected since, even under the null of a constant volatility jump beta, the plots would display noisy estimates of the volatility co-jumps and not the true underlying co-jumps themselves. However, even with this estimation error the co-jumps do generally appear to lie more tightly along the line of best fit than one might expect

if no relationship existed, especially for the portfolios. This is despite the estimation error, the tail nature of the jumps, and the fact that the sample spans both tranquil and turbulent market environments.

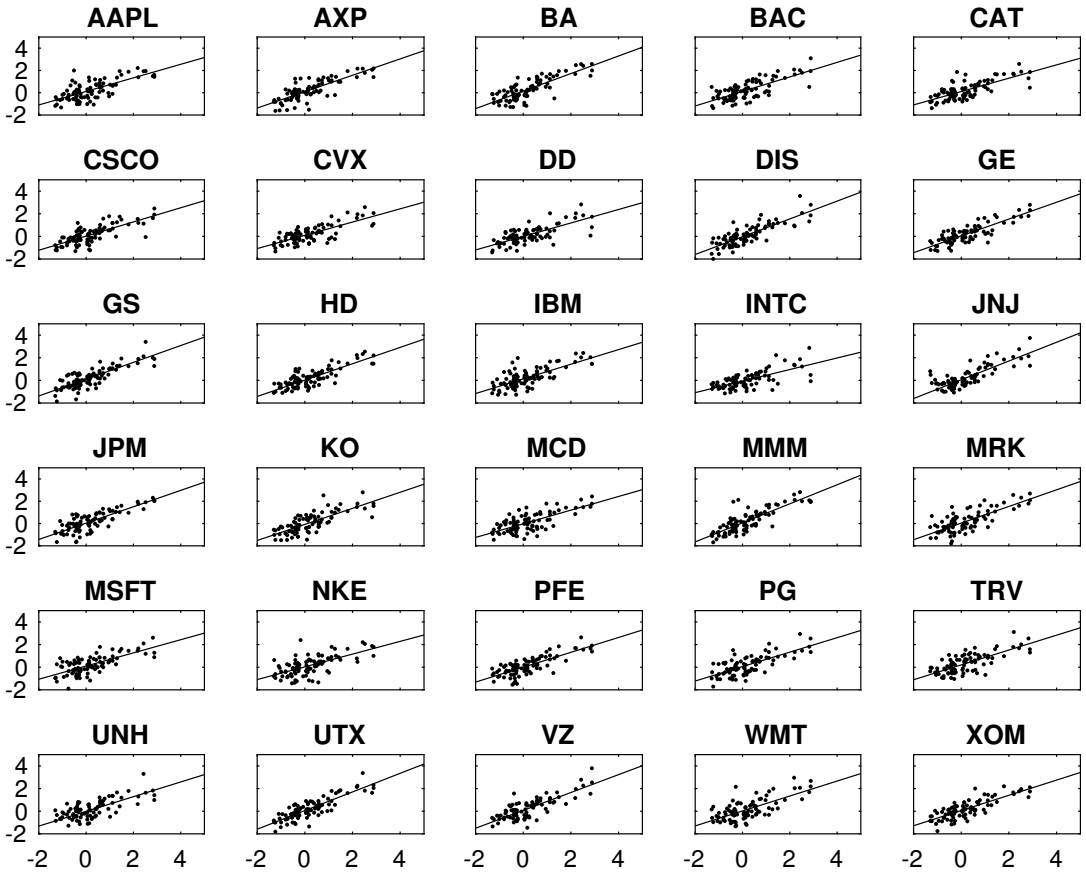
Figure 2: Estimated Volatility Co-Jumps: E-mini vs S&P Industry Portfolios (2007-2015)



Note: This figure displays scatter plots of the estimated volatility co-jumps between the E-mini and the listed asset at estimated jump times in the market for the period 2007-2015. The magnitudes of the volatility co-jumps are estimated following the procedure in Section 3 whereas the market volatility jump times were estimated from the VIX using the procedure detailed in Section 3.2.

Tables 3 and 4 provide summary statistics for the test of a constant volatility jump beta using the determinant test of Section 4.2. The test was conducted for the whole sample and year-by-year. Table 3 provides p-values for the portfolios whereas, to better summarize the results, Table 4 lists the number of stocks in the DOW30 for which one would fail to reject the null of a constant volatility jump beta at significance levels 1%, 5%, and 10%. In general there appears to be little evidence for a constant volatility jump beta across the whole sample. However, analyzed year-by-year the evidence for a constant volatility jump beta is much stronger. Here, with only a few exceptions,

Figure 3: Estimated Volatility Co-Jumps: E-mini vs DOW30 Stocks (2007-2015)



Note: This figure displays scatter plots of the estimated volatility co-jumps between the E-mini and the listed asset at estimated jump times in the market for the period 2007-2015. The magnitudes of the volatility co-jumps are estimated following the procedure in Section 3 whereas the market volatility jump times were estimated from the VIX using the procedure detailed in Section 3.2.

Table 3: Tests for a Constant Volatility Jump Beta (p-values)

Asset	2007	2008	2009	2010	2011	2012	2013	2014	2015	Full Sample
XLB	0.01	0.04	0.44	0.21	0.00	0.00	0.98	0.31	0.19	0.00
XLE	0.06	0.80	0.42	0.00	0.00	0.25	0.45	0.24	0.13	0.00
XLF	0.85	0.00	0.87	0.08	0.00	0.00	0.00	0.51	0.01	0.00
XLI	0.92	0.03	0.47	0.24	0.00	0.22	0.34	0.38	1.00	0.01
XLK	0.03	0.00	0.30	0.03	0.11	0.17	0.68	0.10	0.11	0.00
XLP	0.22	0.18	0.42	0.61	0.00	0.01	0.70	0.16	0.02	0.00
XLU	0.02	0.01	0.21	0.01	0.02	0.00	0.21	0.08	0.01	0.00
XLV	0.01	0.10	0.14	0.04	0.00	0.00	0.28	0.99	0.00	0.00
XLY	0.53	0.00	0.37	0.06	0.07	0.00	0.36	0.28	0.58	0.00
					Number of Estimated Jumps					
	11	13	13	8	9	26	4	4	9	97

Note: This table reports p-values for the test of a constant volatility jump beta using the determinant test of Section 4.2. It does so both year-by-year and for the full sample. The lower section of the table lists the number of estimated market volatility jumps in each year and the entire sample following the details explained in Section 3.2 and implemented following the specifications discussed in Section 6.

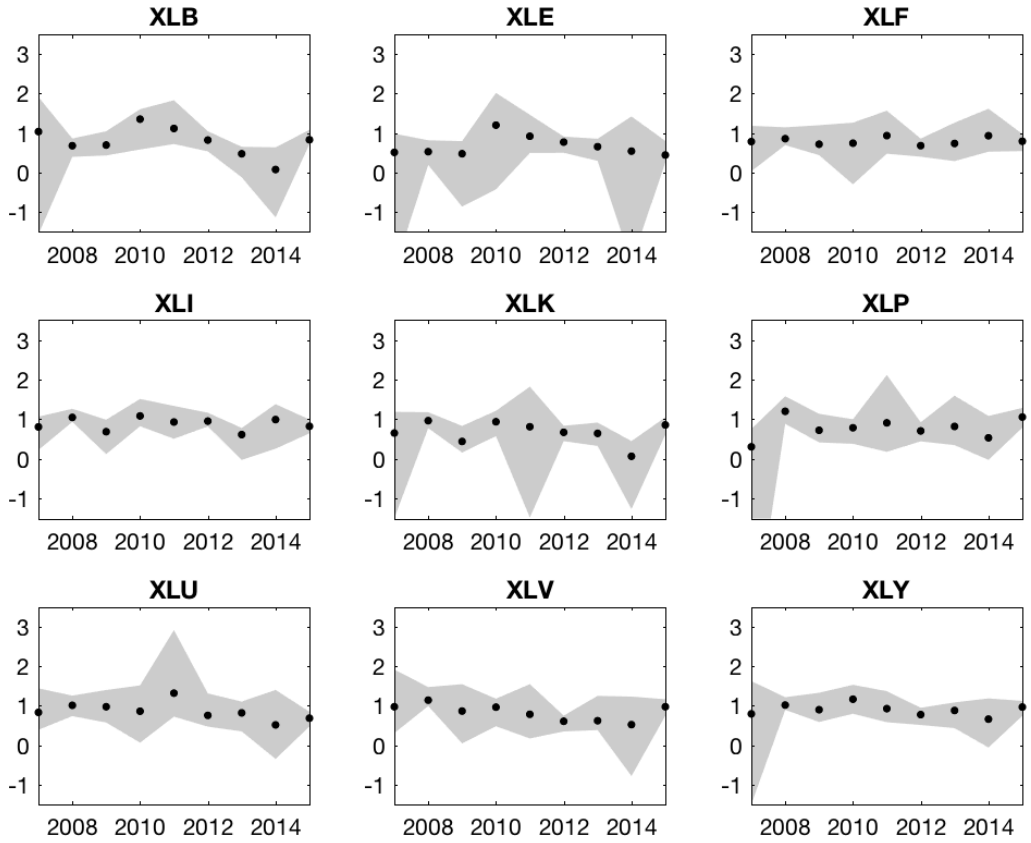
Table 4: Failures to Reject a Constant Volatility Jump Beta for the DOW30 Stocks

Significance Level	2007	2008	2009	2010	2011	2012	2013	2014	2015	Full Sample
1%	13	24	19	24	23	9	23	28	14	0
5%	8	19	13	19	18	4	20	24	6	0
10%	5	18	8	18	12	1	19	19	4	0

Note: The table lists the number of stocks in the DOW30 for which one would fail to reject the null of a constant volatility jump beta at significance levels 1%, 5%, and 10%. The test was done using the determinant test of Section 4.2. The estimation is done both year-by-year and for the full sample.

a constant volatility jump beta cannot be rejected for nearly all of the portfolios at the 1%, 5%, and even 10% level and for a considerable number of the stocks, especially at the 1% level. Such a result might be consistent with variants of a conditional asset pricing model in which the betas change over time (see, for example, Hansen and Richard (1987)).

Figure 4: Time Series Plot of the Volatility Jump Betas for the S&P Industry Portfolios (2007-2015)



Note: This figure plots the estimated volatility jump betas for S&P Industry Portfolios across the years 2007-2015 and their associated 95% confidence intervals. The volatility jump betas are estimated using the quasi-maximum-likelihood method described in Section 4.3 whereas the confidence intervals are calculated using the bootstrap procedure of Section 4.4 and the basic bootstrap using 1000 replications.

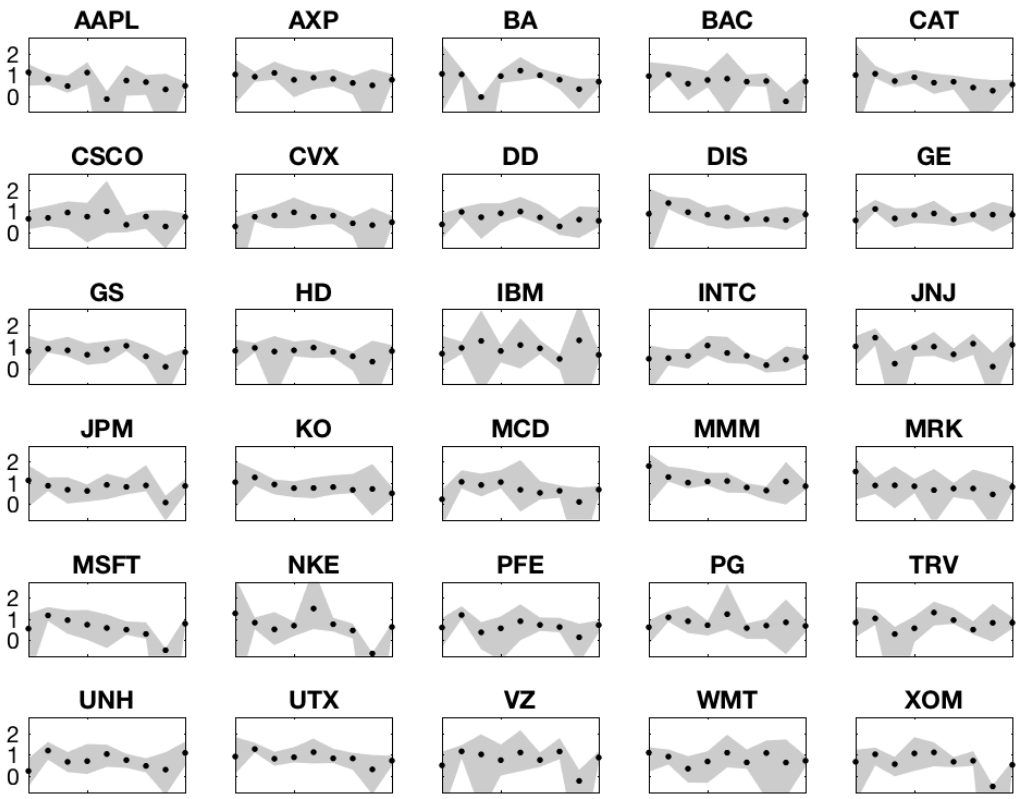
Table 5 lists the estimated volatility jump betas for the portfolios together with their 95% confidence intervals whereas Figures 4 and 5 provide a time series plot of these betas and their confidence intervals. Figure 4 doing so for the portfolios and Figure 5 doing so for the stocks. The volatility jump betas in both the figures and the table were estimated using the quasi-maximum-

Table 5: Estimated Volatility Jump Betas for the S&P Industry Portfolios (2007-2015)

	Full Sample	2007	2008	2009	2010	2011	2012	2013	2014	2015
XLB	0.83	1.04	0.68	0.70	1.35	1.12	0.82	0.48	0.08	0.83
	[0.68, 1.22]	[-1.56, 1.91]	[0.40, 0.86]	[0.43, 1.05]	[0.58, 1.60]	[0.72, 1.83]	[0.54, 1.05]	[-0.12, 0.65]	[-1.13, 0.64]	[0.72, 1.08]
XLE	0.62	0.51	0.53	0.48	1.20	0.92	0.77	0.66	0.54	0.44
	[0.47, 1.02]	[-2.47, 0.98]	[0.19, 0.82]	[-0.86, 0.80]	[-0.42, 2.02]	[0.50, 1.46]	[0.50, 0.91]	[0.30, 0.86]	[-2.27, 1.42]	[0.25, 0.80]
XLF	0.79	0.78	0.86	0.72	0.74	0.94	0.68	0.74	0.93	0.79
	[0.68, 1.03]	[0.03, 1.18]	[0.70, 1.15]	[0.44, 1.20]	[-0.30, 1.26]	[0.48, 1.57]	[0.40, 0.86]	[0.29, 1.26]	[0.53, 1.62]	[0.54, 0.98]
XLI	0.88	0.81	1.05	0.69	1.09	0.93	0.96	0.62	0.99	0.82
	[0.78, 0.95]	[0.23, 1.07]	[0.94, 1.27]	[0.12, 0.99]	[0.83, 1.51]	[0.51, 1.34]	[0.82, 1.17]	[-0.02, 0.78]	[0.26, 1.38]	[0.65, 0.99]
XLK	0.75	0.66	0.97	0.44	0.94	0.81	0.67	0.65	0.07	0.86
	[0.66, 1.11]	[-1.49, 1.19]	[0.78, 1.18]	[0.17, 0.83]	[0.58, 1.21]	[-1.47, 1.82]	[0.45, 0.84]	[0.33, 0.92]	[-1.26, 0.45]	[0.62, 1.05]
XLP	0.86	0.31	1.20	0.72	0.79	0.91	0.71	0.82	0.54	1.06
	[0.60, 1.40]	[-4.74, 0.76]	[0.89, 1.58]	[0.42, 1.14]	[0.38, 1.00]	[0.18, 2.12]	[0.45, 0.93]	[0.35, 1.60]	[-0.02, 1.08]	[0.79, 1.29]
XLU	0.84	0.83	1.01	0.97	0.86	1.32	0.75	0.82	0.51	0.68
	[0.73, 1.09]	[0.39, 1.44]	[0.73, 1.25]	[0.57, 1.39]	[0.06, 1.51]	[0.72, 2.91]	[0.47, 1.31]	[0.35, 1.10]	[-0.35, 1.40]	[0.47, 0.84]
XLV	0.86	0.97	1.14	0.86	0.96	0.78	0.60	0.62	0.52	0.97
	[0.73, 1.00]	[0.31, 1.91]	[0.99, 1.47]	[0.05, 1.54]	[0.48, 1.17]	[0.17, 1.54]	[0.35, 0.75]	[0.38, 1.25]	[-0.78, 1.23]	[0.74, 1.16]
XLY	0.91	0.79	1.01	0.90	1.16	0.92	0.77	0.88	0.66	0.96
	[0.77, 1.00]	[-1.50, 1.62]	[0.89, 1.21]	[0.59, 1.33]	[0.79, 1.53]	[0.58, 1.37]	[0.51, 0.94]	[0.43, 1.08]	[-0.06, 1.18]	[0.74, 1.12]

Note: This table displays the estimated volatility jump betas between the S&P industry portfolios and the E-mini for the entire sample and for each year. Under each point estimate in brackets is a 95% confidence interval. The volatility jump betas were estimated by a quasi-maximum-likelihood procedure detailed in Section 4.3 and the confidence intervals were estimated using the bootstrap procedure detailed in Section 4.4 and the basic bootstrap.

Figure 5: Time Series Plot of the Volatility Jump Betas for the DOW30 Stocks (2007-2015)



Note: This figure plots the estimated volatility jump betas for the DOW30 stocks across the years 2007-2015 and their associated 95% confidence intervals. The volatility jump betas are estimated using the quasi-maximum-likelihood method described in Section 4.3 whereas the confidence intervals are calculated using the bootstrap procedure of Section 4.4 and the basic bootstrap using 1000 replications.

likelihood method described in Section 4.3 and the confidence intervals were calculated using the bootstrap procedure of Section 4.4 and the basic bootstrap with 1000 replications. As noted before, because the limiting bootstrap distribution in Theorem 3 is not symmetric, the confidence intervals will not in general be symmetric.

With a few exceptions the confidence intervals are decently tight, especially if one considers the estimation error involved in estimating the volatility jumps. This shows the fairly good performance of the bootstrap method. For many assets the volatility jump beta does appear to vary from year to year confirming the rejections earlier of a constant volatility jump beta across the sample. Interestingly many of the volatility jump betas appear to be very near, or just slightly below, one. Recall that we are considering changes in the log-volatilities of the assets in this setting. Given this the volatility jump betas correspond to elasticities between the market volatility and the asset volatilities. A log-volatility jump beta of one, or slightly less, would imply that a jump in the market volatility of given percentage would have an equal, or slightly less than equal, percentage change in the volatility of the asset. Such a result would arise if a linear relationship existed between the market volatility jumps and the volatility jumps of the asset. Finding a log-volatility jump beta near one then lends support to modeling asset volatilities as a linear function of the market volatility.

6.3 Idiosyncratic Volatility Jump Regressions

This empirical application estimates and conducts inference for the volatility jump beta between the idiosyncratic volatility of an asset and the market volatility. To illustrate this idea let Z denote the market, or potentially some observed aggregate risk factor, and Y the asset under consideration. We can decompose the continuous parts of Z and Y as follows

$$Y_t^c = \beta_{0,t}^Z + \beta_{1,t}^Z Z_t^c + \widetilde{Y}_t^c \tag{57}$$

where $\widetilde{Y}^c$ is the "idiosyncratic" component in Y^c. If Z^c were the sole risk factor in the economy than $\widetilde{Y}^c$ would be, by construction, an unpriced source of risk. Denoting the volatility of $\widetilde{Y}_t$ as $\widetilde{\sigma}_{Y,t}^2$ and the volatility of Z^c as σ_Z^2, this empirical application considers volatility jump regressions of the form

$$\Delta\widetilde{\sigma}_{Y,\tau_p}^2 = \beta\Delta\sigma_{Z,\tau_p}^2 \tag{58}$$

at the set $(\tau_p)_{p\geq 1}$ of jump times in σ_Z^2. The idiosyncratic volatility is easy to estimate since $\widetilde{\sigma}_{Y,t}^2 = \sigma_{Y,t}^2 - (\sigma_{ZY,t}^2)^2/\sigma_{Z,t}^2$ where $\sigma_{ZY,t}^2$ is the co-volatility between Y^c and Z^c. Note that, as before, the analysis is performed in log units.

Table 6 provides estimates of the idiosyncratic volatility jump betas for the DOW30 stocks and their significance levels. The empirical analysis here only considers the DOW30 stocks since the S&P Sector ETFs are themselves aggregate portfolios. The estimation was conducted for the whole sample and for three consecutive three-year intervals. The significance levels of the estimated betas, that is, whether the betas are estimated to be significantly different from zero are indicated by the stars corresponding to significance levels of 1% (***), 5% (**), or 10% (*). The estimation was performed using the quasi-maximum-likelihood procedure detailed in Section 4.3 and the significance levels were estimated based on the bootstrap confidence intervals detailed in Section 4.4 and the basic bootstrap. Notice the high degree of significance of the estimated betas. With only a few exceptions all the estimated betas are estimated to be significantly different from zero at the 1% level. This provides evidence that, at least at jump times, the idiosyncratic volatility correlates with the market volatility. The R^2s of the regressions are listed in Table 7. Examined across the entire sample the R^2s range from a low of 0.179 for Intel (INTC) to a high of 0.47 for Goldman Sachs (GS) and, examined across the three-year intervals, reach as high as 0.661 for Goldman Sachs during the 2007-2009 period. The high R^2s in this table shows that, not only does the idiosyncratic volatility appear to correlate with the market volatility at jump times, but the market volatility jumps often explains a substantial fraction of the variation in the idiosyncratic volatility jumps themselves.

The magnitudes of the estimated betas are interesting as well. Notice that nearly all of the betas are estimated to be between zero and one and many are in the range 0.4 to 0.7. Recall that we are considering changes in the log-volatilities of the assets in this setting. Given this the volatility jump betas correspond to elasticities between the market volatility and the idiosyncratic volatilities of the assets. A beta less than one then implies a concave relationship between the market volatility jumps and the idiosyncratic volatility jumps.

The results of this empirical application shed light on the growing literature on idiosyncratic risk and the so-called "idiosyncratic volatility puzzle." The "puzzle" was first documented in Ang, Hodrick, Xing, and Zhang (2006) and concerns the fact that the idiosyncratic volatility of assets appears to be priced in the cross-section. Ang, Hodrick, Xing, and Zhang (2006) used changes in the VIX to proxy for changes in the market volatility and calculated the idiosyncratic volatility based on the residuals from regressions of asset returns on the changes in the VIX together with the three factors in Fama and French (1993). Sorting stocks based on their average idiosyncratic volatility they found that stocks with higher idiosyncratic volatilities had lower average returns and analogously stocks with lower idiosyncratic volatilities had higher average returns. They explored

Table 6: Estimated Idiosyncratic Volatility Jump Betas for the DOW30 Stocks (2007-2015)

Asset	Full Sample	2007-2009	2010-2012	2013-2015
AAPL	0.620***	0.951***	0.385*	0.367***
AXP	0.628***	0.900***	0.280*	0.533***
BA	0.622***	0.760***	0.727***	0.348***
BAC	0.708***	0.830***	0.453***	0.792***
CAT	0.543***	0.865***	0.507***	0.258***
CSCO	0.459***	0.654***	0.109	0.484***
CVX	0.389***	0.586***	0.326**	0.262
DD	0.515***	0.696***	0.433***	0.417**
DIS	0.714***	0.989***	0.337***	0.701***
GE	0.597***	0.785***	0.352**	0.614***
GS	0.643***	0.804***	0.619***	0.489***
HD	0.735***	0.986***	0.479***	0.717***
IBM	0.575***	0.798***	0.470***	0.455***
INTC	0.292***	0.335***	0.412***	0.150
JNJ	0.758***	0.872***	0.504***	0.916***
JPM	0.734***	0.874***	0.369**	0.784***
KO	0.596***	0.836***	0.447***	0.449***
MCD	0.553***	0.886***	0.376**	0.384***
MMM	0.766***	1.146***	0.575***	0.595***
MRK	0.632***	0.902***	0.531***	0.461***
MSFT	0.487***	0.759***	0.405**	0.299**
NKE	0.540***	0.768***	0.502***	0.387***
PFE	0.502***	0.633***	0.397***	0.463***
PG	0.606***	0.855***	0.458***	0.466***
TRV	0.602***	0.732***	0.551**	0.531***
UNH	0.668***	0.620***	0.683***	0.699***
UTX	0.660***	0.937***	0.404***	0.564***
VZ	0.687***	0.780***	0.507**	0.739***
WMT	0.752***	0.965***	0.463***	0.755***
XOM	0.492***	0.711***	0.426***	0.311**

Note: This table displays the estimated idiosyncratic volatility jump betas between the DOW30 stocks and the E-mini for the entire sample and for three consecutive three-year intervals. The stars are indications that the estimated beta is significantly different from zero at either the 1% (***), 5% (**), or 10% (*) level. The volatility jump betas were estimated by a quasi-maximum-likelihood procedure detailed in Section 4.3 and the significance levels were estimated based on the bootstrap confidence intervals detailed in Section 4.4 and the basic bootstrap.

Table 7: Idiosyncratic Volatility Jump Regression R^2s for the DOW30 Stocks (2007-2015)

Asset	Full Sample	2007-2009	2010-2012	2013-2015
AAPL	0.233	0.589	0.009	0.315
AXP	0.317	0.601	0.018	0.605
BA	0.383	0.475	0.470	0.252
BAC	0.405	0.595	0.245	0.383
CAT	0.277	0.467	0.148	0.371
CSCO	0.288	0.549	0.012	0.542
CVX	0.231	0.356	0.194	0.166
DD	0.241	0.317	0.271	0.161
DIS	0.361	0.493	0.132	0.507
GE	0.348	0.488	0.195	0.373
GS	0.470	0.661	0.331	0.472
HD	0.400	0.497	0.388	0.348
IBM	0.286	0.380	0.112	0.469
INTC	0.179	0.135	0.316	0.120
JNJ	0.415	0.415	0.502	0.359
JPM	0.325	0.512	0.028	0.548
KO	0.346	0.571	0.223	0.288
MCD	0.228	0.316	0.108	0.331
MMM	0.411	0.404	0.420	0.504
MRK	0.289	0.343	0.242	0.324
MSFT	0.187	0.353	0.051	0.266
NKE	0.269	0.260	0.226	0.370
PFE	0.362	0.497	0.197	0.432
PG	0.252	0.433	0.109	0.286
TRV	0.306	0.328	0.139	0.520
UNH	0.190	0.290	0.081	0.214
UTX	0.444	0.659	0.292	0.411
VZ	0.353	0.452	0.194	0.423
WMT	0.263	0.428	0.178	0.204
XOM	0.304	0.534	0.158	0.284

Note: This table displays the estimated idiosyncratic volatility jump betas between the DOW30 stocks and the E-mini for the entire sample and for three consecutive three-year intervals. The stars are indications that the estimated beta is significantly different from zero at either the 1% (***), 5% (**), or 10% (*) level. The volatility jump betas were estimated by a quasi-maximum-likelihood procedure detailed in Section 4.3 and the significance levels were estimated based on the bootstrap confidence intervals detailed in Section 4.4 and the basic bootstrap.

the possibility that such a finding might be due to fact that stocks with high idiosyncratic volatilities may have a high exposure to market volatility risk, which would be in line with the findings here. They however found this to be an incomplete explanation. In particular, they found that stocks that high (low) idiosyncratic volatilities had a high (low) loading on market volatility risk, but that this relationship only existed for stocks with low previous loadings on the market volatility. One potentially issue with their results is their use of changes in the VIX to proxy for changes in the market volatility. Since the VIX is the expectation of the future implied volatility of the market under the risk-neutral density, using the VIX in place of measures of volatility based on realized returns conflates changes in the market volatility with potential changes in the volatility risk premium. The results here, while obviously limited to volatility jump events, do not have this problem since the estimates are based on realized high-frequency returns. Many of the explanations for the negative price of idiosyncratic risk following Ang, Hodrick, Xing, and Zhang (2006) have involved positing lottery-like preferences for investors or have examined market fractions. For a thorough overview of these explanations and the literature on idiosyncratic risk since Ang, Hodrick, Xing, and Zhang (2006) see Hou and Loh (2016).

6.4 The Pricing of Volatility Jumps

In addition to providing a method for estimating volatility jump betas, the procedures in this paper provide a means to estimate the timing and magnitude of jumps in the market volatility. Such methods can be used to explore the pricing of volatility risk. To do so this study performs a series of predictive regressions between the estimated market volatility jumps and future excess returns. To examine these regressions let $er_{t,t+h}$ be the excess return of some asset between days t and $t+h$ and let t_p be the day on which the p-th market volatility jump takes place. The regressions take the form

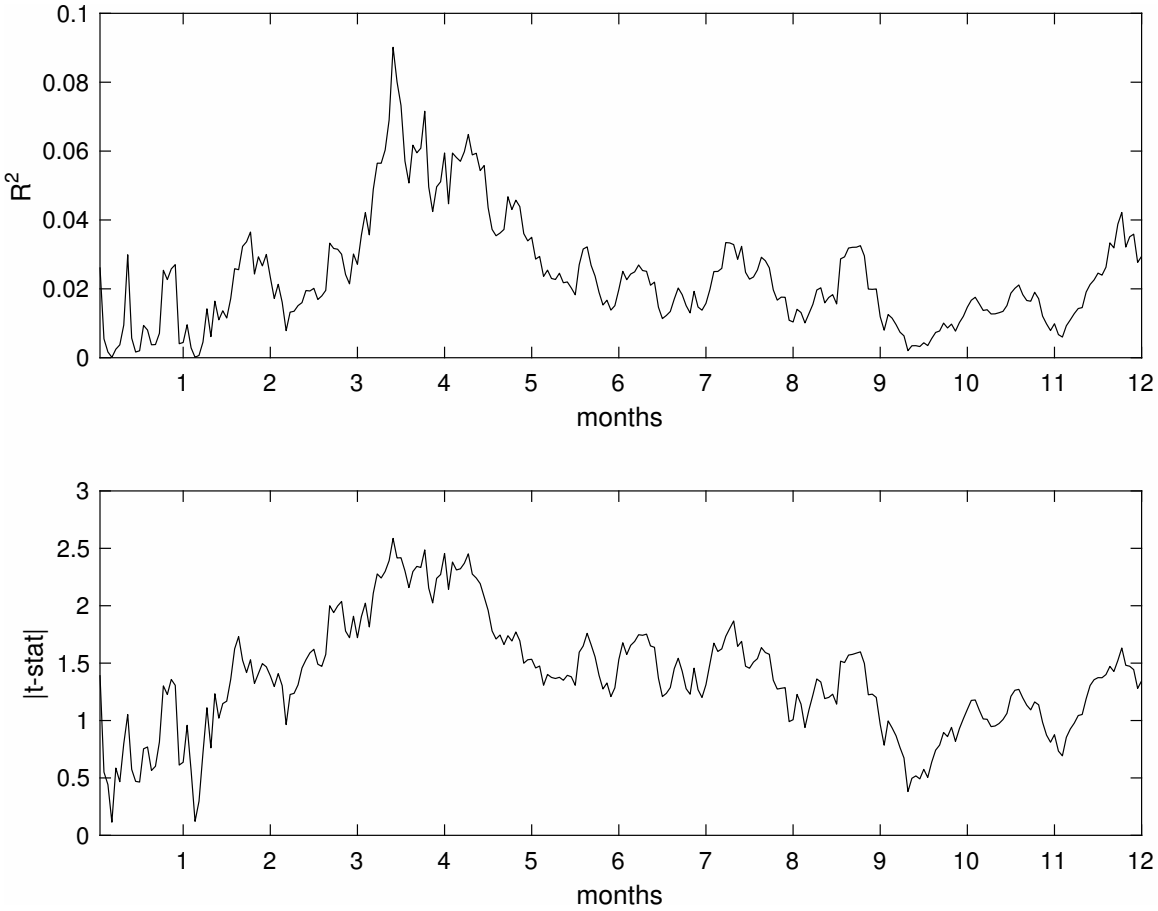
$$er_{t_p+1,t_p+h} = \gamma_0 + \gamma_1 \Delta\hat{\sigma}^2_{m,i_p} + \epsilon_{t_p} \tag{59}$$

for all $p \geq 1$ where $\hat{\sigma}^2_{m,i_p}$ is the estimated p-th market volatility jump.

Figures 6 and 7 plot the R^2s and t-statistics from these predictive regressions. Figure 6 doing so for the predictive regressions of the excess returns on the market, with the E-mini being the proxy for the market, and Figure 7 doing so for the excess returns on the DOW30 stocks. The t-statistics were calculated using the heteroskedastic-robust method in White (1980).[6] Examining Figure 6

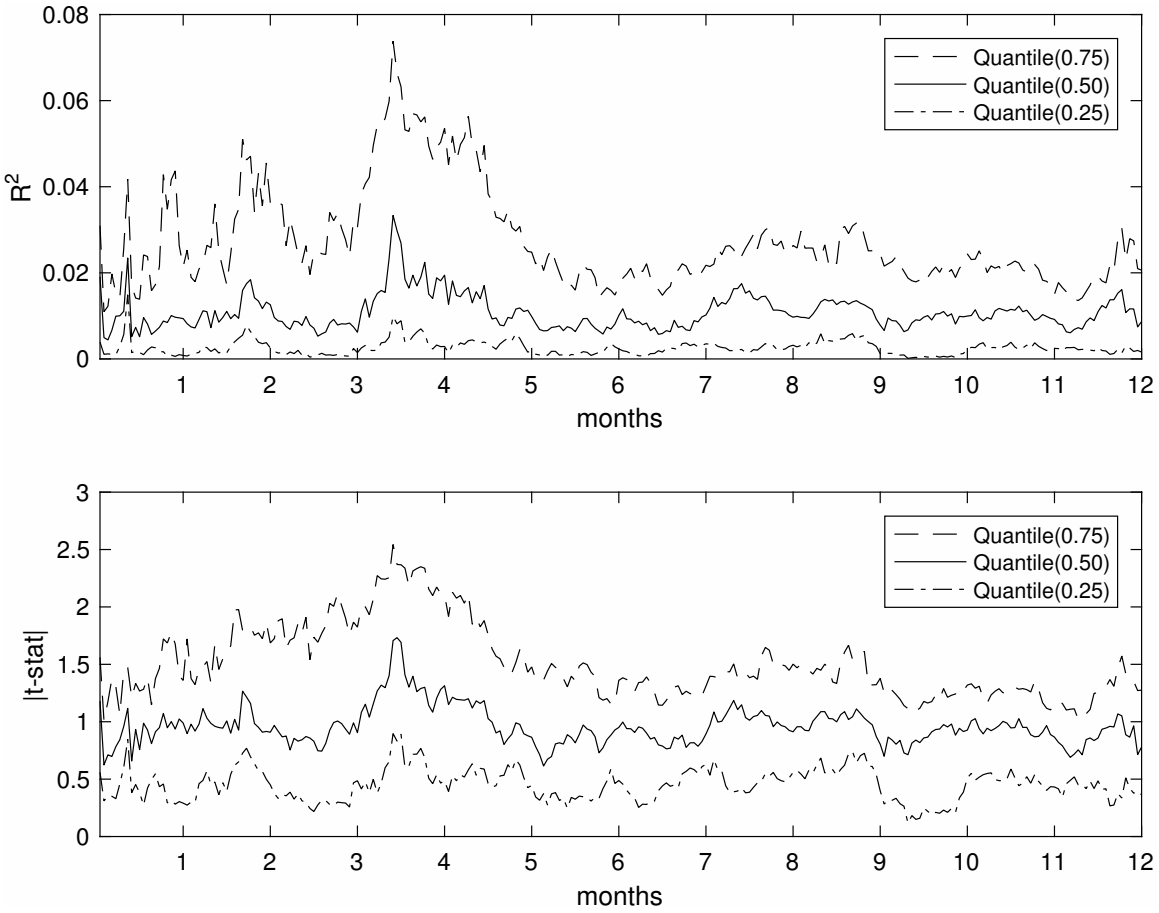
[6]Note that due to the stochastic arrival of the jump times that any serial correlation in the process is greatly reduced and would not be dealt with correctly by standard time-series methods. A more detailed study might consider appropriate methods to handle the error structure in these predictive regressions. Also note that these regressions require the use of both in-fill and long-span asymptotics. The in-fill asymptotics to estimate the spot volatility jumps and the long-span asymptotics to justify the convergence properties of the regressions.

Figure 6: Predictive Regressions: Market Volatility Jumps and the Market Excess Returns



Note: This figure plots predictive regression results for predictive regressions between the estimated market volatility jumps and the market excess returns using the E-mini as the proxy for the market and the three-month T-bill rate as the proxy for the risk-free rate. The sample spans the years 2007-2015. The top panel plots the R^2s from these regressions and the bottom panel the absolute value of the t-statistic. The t-statistics were calculated using the heteroskedastic robust procedure in White (1980).

Figure 7: Predictive Regressions: Market Volatility Jumps and the DOW30 Excess Returns



Note: This figure plots predictive regression results for predictive regressions between the estimated market volatility jumps and the excess returns of the DOW30 stocks where the three-month T-bill rate is used as the proxy for the risk-free rate. The sample spans the years 2007-2015. The top panel plots the 75th, 50th, and 25th quantiles of the R^2s from these regressions and the bottom panel plots the corresponding absolute t-statistics. The t-statistics were calculated using the heteroskedastic robust procedure in White (1980).

notice the high degree of predictability that the market volatility jumps have on future market excess returns especially at a three to five month horizon. The series peaks at the 75 day horizon excess return with an $R^2 = 0.09$. Figure 7 provides corresponding results for the DOW30 stocks. The figure plots quantiles of the R^2s and t-statistics across the thirty stocks for these predictive regressions. While not as large in magnitude as the results for the market excess return, the figure shows that market volatility jumps remain highly predictive for future excess stock returns. These result are similar to the findings of other studies. Bollerslev, Tauchen, and Zhou (2009), for example, found the variance risk premium to be most predictive for future excess returns at a three to six month horizon and found an R^2 of about 0.07 at most. Whereas, Andersen, Fusari, and Todorov (2015c) found their left-jump tail-risk factor to be most predict for future excess returns at about a six month horizon found R^2s of no more than about 0.14.

A guide to understand the pricing of volatility jumps follows from the recent work of Andersen, Fusari, and Todorov (2015a) and Andersen, Fusari, and Todorov (2015c). To do so we will need to introduce a risk-neutral dynamic for the log-price process X_t under the risk-neutral measure $\mathbb{Q}$. Recall the dynamics of X_t under the physical measure $\mathbb{P}$ as originally defined in (3). To specify the dynamics of X_t under $\mathbb{Q}$ define

$$W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} + \int_0^t \lambda_s^W ds \tag{60}$$

where $W^{\mathbb{Q}}$ is a Brownian motion under $\mathbb{Q}$ and λ_s^W is the price of risk due to W, i.e., the price of diffusive risk. The Poisson measure under $\mathbb{Q}$ is given by $\underline{p}^{\mathbb{Q}}$ with associated compensator $\underline{q}^{\mathbb{Q}}(dt, dx) = dt \otimes \lambda^{\mathbb{Q}}(dx)$ where the mapping $\lambda^{\mathbb{Q}}(dx) \to \lambda^{\mathbb{P}}(dx)$ reflects the compensation for jump risk. With these definitions a risk-neutral dynamic for X_t exists with the following structure

$$\begin{aligned} dX_t &= (r_t - \delta_t)dt + \sigma_t dW_t^{\mathbb{Q}} \\ &\quad + \int_{\mathbb{R}} \left(\delta(x) 1_{\{||\delta(x)|| \leq 1\}}\right) * \left(\underline{p}^{\mathbb{Q}} - \underline{q}^{\mathbb{Q}}\right)(dt, dx) + \left(\delta(x) 1_{\{||\delta(x)|| > 1\}}\right) * \underline{p}^{\mathbb{Q}}(dt, dx) \end{aligned} \tag{61}$$

where σ_t is the same volatility process as in (3), r_t is the risk-free rate, and δ_t is the dividend-yield. Recall that under $\mathbb{P}$ the drift in X_t was denoted by b_t. With these definitions the spot equity risk premium is given by

$$b_t - (r_t - \delta_t) = \lambda_t^W \sigma_t + \int_{\mathbb{R}} \delta(x)(\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}})(dx) \tag{62}$$

where the two components reflects compensation for diffusive risk and jump risk respectively. Notice how volatility jumps, that is jumps in σ_t, enter into the equity risk premium (and by extension into the drift of the log-price process). This shows that future excess returns will be affected by volatility jumps through their cumulative impact on the path of the spot equity risk premium process.

7 Conclusion

This paper develops econometric tools for studying the jump dependencies between functions of the underlying or latent spot co-volatilities of two assets from high-frequency observations on a fixed time interval with a particular interest on the relationship between functions of the spot volatility of different traded assets and the spot volatility of aggregate risk factors such as the market volatility. This paper derives a test for deciding whether a linear relationship between functions of the latent spot volatility jumps of two assets exists on a given time interval which, as explained in the paper, could have potentially very interesting implications for the modeling or pricing of volatility processes. The test is shown to have power for detecting both nonlinearity in functional form and time-varying parameters. Next, after providing evidence that an ordinary least squares estimator for the volatility jump beta will be biased in finite samples, the paper proposes a quasi-maximum-likelihood based estimator to correct for this bias and shows it to be asymptoticly valid and consistent for the volatility jump beta. The quasi-maximum-likelihood estimator is based on an asymptotic approximation of the error distribution in estimating the volatility jump relationship. Because the limiting distribution of the estimator is non-standard a simple and intuitive bootstrap method is provided and shown to be valid for feasible inference. The bootstrap is based on a simple i.i.d. re-sampling of the returns in a neighborhood around the volatility jump times. In addition, the paper also shows how a volatility index, such as the Chicago Board of Option Exchange's (CBOE) Volatility Index (VIX), can be used to identify the latent jump times in the market volatility under fairly minor assumptions. This addresses the issue of locating the spot volatility jump times. In a set of Monte Carlo experiments over a range of parametric specifications and sampling frequencies the paper reports good size and power properties for the specification test and good coverage properties for confidence intervals. In addition, the study shows the proposed quasi-maximum-likelihood estimator to have a markedly smaller bias than the ordinary least squares estimator in these trials.

The paper contains three empirical applications. The first study employs a panel consisting of the E-mini S&P500 index futures, the nine S&P portfolio ETFs, and the thirty stocks in the Dow Jones Industrial Average over the period 2007-2015 with the E-mini being the proxy for the market. Comparing log-volatility jumps the study documents that the market volatility jump beta of many financial assets appears to remain stable over a period of one year, but finds evidence for temporal variation over the entire sample. Interestingly, the volatility jump beta for many assets appears to be very near one which, given the comparison between log-volatilities, would imply a unit elastic relationship between the volatility of the market and the volatility of the assets at jump times in

the market volatility. This would in turn would imply a linear relationship between the volatility of these assets and the market volatility at jump times. The next study examines the beta between the "idiosyncratic" volatilities of assets and the market volatility. The study finds many of these betas to be statistically different from zero implying a correlation between the idiosyncratic volatilities of these assets and the market volatility at jump times. Even further, the R^2s of these regressions are often in excess of 0.50 implying that, at jump times, the market volatility often explains a substantial fraction of the variation in the idiosyncratic volatilities. The final empirical application examines the pricing of market volatility jumps. Through a series of predictive regressions between future market excess returns and the estimated market spot volatility jumps, the study documents market spot volatility jumps to be particularly predictive of future excess returns at a three to four month horizon with R^2s of nearly ten percent.

8 Appendix: Proofs

We now prove the theorems in the main text. Throughout this appendix, we use K to denote a generic positive constant that may change from line to line; we sometimes emphasize the dependence of this constant on some parameter q by writing K_q. Below, the convergence $(\xi_{n,p})_{p\geq 1} \to (\xi_p)_{p\geq 1}$ as $n \to \infty$, is understood to be under the product topology. We use the following sets of notation. First, we write w.p.a.1 for "with probability approaching one." Next we define $\mathcal{T} = (\tau_p)_{p\geq 1}$ where $(\tau_p)_{p\geq 1}$ are the successive jump times of $c_{ZZ,t}$ and denote $\mathcal{P} = \{p \geq 1 : \tau_p \in [0,T]\}$. Finally, we use i_p to denote the unique integer such that $\tau_p \in ((i-1)\Delta_n, i\Delta_n]$.

By a standard localization procedure (see Section 4.4.1 of Jacod and Protter (2012)), we can respectively strengthen Assumptions 1 and 2 to the following stronger versions without loss of generality.

Assumption 4. *We have Assumption 1. Moreover, the processes X_t, b_t, and σ_t are bounded.*

Assumption 5. *We have Assumption 2. Moreover, the processes $b_t^{(c)}$ and $\sigma_t^{(c)}$ are bounded.*

Proof of Proposition 1. Part (a) follows by a simple application of Proposition 1 in Li, Todorov, and Tauchen (2015). Part (b) follows by verifying the assumptions in Theorem 1 in Todorov and Tauchen (2011). □

Proof of Theorem 1. By Proposition 1 we have that

$$\mathbb{P}(\mathscr{I}_n = \mathscr{I}_n^*) \to 1. \tag{63}$$

Then, by Theorem 9.3.2 in Jacod and Protter (2012) we derive that

$$\begin{aligned}
&\big(g(\hat{c}_{i_{p-}})\,,\, g(\hat{c}_{i_{p+}}), h(\hat{c}_{i_{p-}}), h(\hat{c}_{i_{p+}})\big)_{i_p \in \mathscr{I}_n} \\
&\qquad = \big(g(\hat{c}_{i_{p-}}), g(\hat{c}_{i_{p+}}), h(\hat{c}_{i_{p-}}), h(\hat{c}_{i_{p+}})\big)_{i_p \in \mathscr{I}_n^*} \text{ w.p.a.1 by (63)} \\
&\qquad \xrightarrow{\mathbb{P}} \big(g(c_{\tau_{p-}}), g(c_{\tau_{p+}}), h(c_{\tau_{p-}}), h(c_{\tau_{p+}})\big)_{p\in\mathcal{P}}\,.
\end{aligned} \tag{64}$$

Recall the notation in Sections 3 and 4. By Theorem 8.8 in Aït-Sahalia and Jacod (2014) and the delta method we have

$$\begin{aligned}
&\sqrt{k_n}\left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p}), \Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p})\right)_{i_p \in \mathscr{I}_n} \\
&\quad = \sqrt{k_n}\left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p}), \Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p})\right)_{i_p \in \mathscr{I}_n^*} \text{ w.p.a.1 by (63)} \\
&\quad \xrightarrow{\mathcal{L}-s} \left(\nabla g(c_{\tau_p})^\top U_{\tau_p} + \nabla g(c_{\tau_{p-}})^\top U_{\tau_{p-}}, \nabla h(c_{\tau_p})^\top U_{\tau_p} + \nabla h(c_{\tau_{p-}})^\top U_{\tau_{p-}}\right)_{p\in\mathcal{P}} \\
&\quad \equiv \big(\xi_{\tau_p}(g), \xi_{\tau_p}(h)\big)_{p\in\mathcal{P}}\,.
\end{aligned} \tag{65}$$

where the variables $U_{q-},\ U_q,\ U'_{q-}$, and U'_q are defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and, conditionally on $\mathcal{F}$, are all independent centered Gaussian with covariances given by (10).

(a) Define

$$\bar{\epsilon}_{i_p}(g,h) = \big(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p})\big) - \bar{\beta}(g,h)\left(\Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p})\right) \tag{66}$$

and note that in restriction to Ω_0 we have

$$\Delta g(\hat{c}_{i_p}) = \beta_0(g,h)\Delta h(\hat{c}_{i_p}) + \bar{\epsilon}_{i_p}(g,h) \tag{67}$$

since $\bar{\beta}(g,h) = \beta_0(g,h)$ in restriction to Ω_0. Consequently by (65) we have

$$\sqrt{k_n}\bar{\epsilon}_{i_p}(g,h) \xrightarrow{\mathcal{L}-s} \bar{\varphi}_{\tau_p}(g,h) \tag{68}$$

for all $p \geq 1$ where $\bar{\varphi}_{\tau_p}(g,h) = \xi_{\tau_p}(g) - \bar{\beta}(g,h)\xi_{\tau_p}(h)$.

Because $\mathbb{P}(\mathscr{I}_n = \mathscr{I}_n^*) \to 1$ w.p.a.1,

$$\det[Q_n(g,h)] = \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})^2\right)\left(\sum_{i_p \in \mathscr{I}_n^*} \Delta g(\hat{c}_{i_p})^2\right) - \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})\Delta g(\hat{c}_{i_p})\right)^2. \tag{69}$$

Plugging (66) into (69) we see, after some algebra, that

$$\det[Q_n(g,h)] = \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})^2\right)\left(\sum_{i_p \in \mathscr{I}_n^*} \bar{\epsilon}_{i_p}^2(g,h)\right) - \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})\bar{\epsilon}_{i_p}(g,h)\right)^2. \tag{70}$$

Combining the convergence in (64) with the convergence in (68) we use the property of stable convergence to derive the joint convergence

$$\big(\Delta h(\hat{c}_{i_p}), \bar{\epsilon}_{i_p}(g,h)\big)_{p\geq 1} \xrightarrow{\mathcal{L}-s} \big(\Delta h(c_{\tau_p}), \bar{\varphi}_{\tau_p}(g,h)\big)_{p\geq 1}. \tag{71}$$

Since the set $\mathcal{P}$ is a.s. finite, the assertion of part (a) follows from (70), (71), and the continuous mapping theorem.

(b) By a standard localization argument (see Section 4.4.1 of Jacod and Protter (2012)), we can assume that Assumptions 4 and 5 hold without loss of generality. Since $\mathcal{P}$ is a.s. finite, we can also assume that $|\mathcal{P}| \leq M$ for some constant $M > 0$ for the purpose of proving convergence in probability; otherwise, we can fix some large M to make $\mathbb{P}(|\mathcal{P}| > M)$ arbitrarily small and restrict the calculation below on the set $\{|\mathcal{P}| \leq M\}$.

By (63) we have

$$Q_n(g,h) \xrightarrow{\mathbb{P}} Q(g,h) \tag{72}$$

which implies

$$\bar{\beta}_n(g,h) \xrightarrow{\mathbb{P}} \bar{\beta}(g,h). \tag{73}$$

Together with (63) these imply

$$\tilde{\zeta}_n(g,h) = \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})^2\right)\left(\sum_{i_p \in \mathscr{I}_n^*} \tilde{\varphi}_{i_p}^2(g,h)\right) - \left(\sum_{i_p \in \mathscr{I}_n^*} \Delta h(\hat{c}_{i_p})\tilde{\varphi}_{i_p}(g,h)\right)^2 \text{ w.p.a.1.} \tag{74}$$

Next, fix any subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$. By (65) and (73), we can extract a further subsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$, such that along $\mathbb{N}_2$,

$$\left(\left(\hat{c}_{i_{p-}}, \hat{c}_{i_{p+}}\right)_{1\leq p\leq M}, \bar{\beta}_n(g,h)\right) \to \left(\left(c_{\tau_{p-}}, c_{\tau_{p+}}\right)_{1\leq p\leq M}, \beta_0(g,h)\right) \tag{75}$$

on some set Ω^* with $\mathbb{P}(\Omega^*) = 1$. Then, for each $\omega \in \Omega^*$ fixed, the transition kernel of $\tilde{\zeta}_n(g,h)$ given $\mathcal{F}$ converges weakly to the $\mathcal{F}$-conditional law of $\zeta(g,h)$. Moreover, observe that the $\mathcal{F}$-conditional law of the variables $(\bar{\varphi}_{\tau_p}(g,h))_{1\leq p\leq M}$ does not have atoms and has full support on $\mathbb{R}^M$. Therefore, the $\mathcal{F}$-conditional distribution function of $\zeta(g,h)$ is continuous and strictly increasing. By Lemma 21.2 in van der Vaart (1998), we deduce that on each $\omega \in \Omega^*$ and along the subsequence $\mathbb{N}_2$ that $cv_n^\alpha \to cv^\alpha$, where cv^α is the $\mathcal{F}$-conditional $(1-\alpha)$-quantile of $\zeta(g,h)$. Since the subsequence $\mathbb{N}_1$ was arbitrarily chosen, we further deduce that $cv_n^\alpha \to cv^\alpha$ by the subsequence characterization of convergence in probability. This concludes the proof of (b).

(c) By parts (a) and (b), as well as the property of stable convergence, we have

$$\left(k_n \det[Q_n(g,h)], cv_n^\alpha, \mathbb{1}(\Omega_0)\right) \xrightarrow{\mathcal{L}-s} \left(\zeta(g,h), cv^\alpha, \mathbb{1}(\Omega_0)\right). \tag{76}$$

In particular,

$$\mathbb{P}\left(\{k_n \det[Q_n(g,h)] > cv_n^\alpha\} \cap \Omega_0\right) \longrightarrow \mathbb{P}\left(\{\zeta(g,h) > cv^\alpha\} \cap \Omega_0\right). \tag{77}$$

Since $\mathbb{P}(\zeta(g,h) > cv^\alpha|\mathcal{F}) = \alpha$ and $\Omega_0 \in \mathcal{F}$, we have

$$\mathbb{P}\left(\{\zeta(g,h) > cv^\alpha\} \cap \Omega_0\right) = \alpha\mathbb{P}(\Omega_0). \tag{78}$$

The first assertion of part (c) then follows from (77). To show the second assertion of part (c), we first observe that (72) implies $\det[Q_n(g,h)] \xrightarrow{\mathbb{P}} \det[Q(g,h)]$. In restriction to Ω_a we have that $\det[Q(g,h)] > 0$ which implies that $k_n \det[Q_n(g,h)]$ diverges to $+\infty$ in probability. Part (b) implies that cv_n^α is tight in restriction to Ω_a. Consequently, $\mathbb{P}\left(k_n \det[Q_n(g,h)] > cv_n^\alpha|\Omega_a\right) \to 1$ as asserted. □

Proof of Theorem 2. Before proceeding to the proofs of parts (a) and (b) we provide some initial results. First by Proposition 1 we have that

$$\mathbb{P}(\mathscr{I}_n = \mathscr{I}_n^*) \to 1. \tag{79}$$

Next, we formalize some notation and results in the text. Let $\log \mathcal{L}$ and $\log \hat{\mathcal{L}}_n$ be defined as (43) and (44). Next, let U_{q-} and U_q be defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and, conditionally on $\mathcal{F}$, all independent centered Gaussian with covariances

$$\begin{aligned} \widetilde{\mathbb{E}}[U_{q-}^{ij} U_{q-}^{kl}] &= c_{S_{q-}}^{ik} c_{S_{q-}}^{jl} + c_{S_{q-}}^{il} c_{S_{q-}}^{jk}, \\ \widetilde{\mathbb{E}}[U_{q}^{ij} U_{q}^{kl}] &= c_{S_{q}}^{ik} c_{S_{q}}^{jl} + c_{S_{q}}^{il} c_{S_{q}}^{jk}. \end{aligned} \tag{80}$$

and define

$$\begin{aligned} \xi_{\tau_p}(g) &= \nabla g(c_{\tau_p})^\top U_{\tau_p} + \nabla g(c_{\tau_{p-}})^\top U_{\tau_{p-}} \\ \xi_{\tau_p}(h) &= \nabla h(c_{\tau_p})^\top U_{\tau_p} + \nabla h(c_{\tau_{p-}})^\top U_{\tau_{p-}} \end{aligned} \tag{81}$$

By Theorem 8.8 in Aït-Sahalia and Jacod (2014), together with the delta method, we have that

$$\sqrt{k_n} \left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p}), \Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p}) \right) \xrightarrow{\mathcal{L}-s} \left(\xi_{\tau_p}(g), \xi_{\tau_p}(h) \right) \tag{82}$$

for all $p \geq 1$ as $k_n \to \infty$ and $k_n \sqrt{\Delta_n} \to 0$. For all $p \geq 1$, let the covariance matrix of $\left(\xi_{\tau_p}(g), \xi_{\tau_p}(h) \right)$ be defined as $\Sigma_{\tau_p} = \Sigma(g, h, c_{\tau_{p-}}, c_{\tau_{p+}})$ and define its finite sample approximation as $\hat{\Sigma}_{i_p} = \Sigma(g, h, \hat{c}_{i_{p-}}, \hat{c}_{i_{p+}})$. By Theorem 8.6 in Aït-Sahalia and Jacod (2014) together with the continuous mapping theorem and the result in (79) we have

$$\begin{aligned} \sqrt{k_n} \left(\hat{\Sigma}_{i_p} \right)_{i_p \in \mathscr{I}_n} &= \sqrt{k_n} \left(\hat{\Sigma}_{i_p} \right)_{i_p \in \mathscr{I}_n^*} \text{ w.p.a.1 by (79)} \\ &\xrightarrow{\mathbb{P}} \left(\Sigma_{\tau_p} \right)_{p \in \mathcal{P}}. \end{aligned} \tag{83}$$

(a) The proof proceeds by verifying the conditions of Theorem 2.7 in Newey and McFadden (1994). First, by the information inequality it is easy to verify that $\beta_0(g, h)$ uniquely maximizes $\log \mathcal{L}$. The concavity of $\log \hat{\mathcal{L}}_n$ is also readily verifiable. We proceed then by showing the convergence in probability of $\log \hat{\mathcal{L}}_n$ to $\log \mathcal{L}$.

Let

$$\varphi_{\tau_p}(g, h) = \xi_{\tau_p}(g) - \beta_0(g, h) \xi_{\tau_p}(h) \tag{84}$$

and

$$\epsilon_{i_p}(g, h) = \left(\Delta g(\hat{c}_{i_p}) - \Delta g(c_{\tau_p}) \right) - \beta_0(g, h) \left(\Delta h(\hat{c}_{i_p}) - \Delta h(c_{\tau_p}) \right). \tag{85}$$

for all $p \geq 1$ as in the text. By (79), (82), and the continuous mapping theorem note that

$$\begin{aligned} \sqrt{k_n} \left(\epsilon_{i_p}(g, h) \right)_{i_p \in \mathscr{I}_n} &= \sqrt{k_n} \left(\epsilon_{i_p}(g, h) \right)_{i_p \in \mathscr{I}_n^*} \text{ w.p.a.1 by (79)} \\ &\xrightarrow{\mathcal{L}-s} \left(\varphi_{\tau_p}(g, h) \right)_{p \in \mathcal{P}}. \end{aligned} \tag{86}$$

This implies

$$\epsilon_{i_p}(g,h) = k_n^{-1/2}\varphi_{\tau_p}(g,h) + o_p(1) \tag{87}$$

for all $p \geq 1$. The log-likelihood of $\left(k_n^{-1/2}\varphi_{\tau_p}(g,h)\right)_{p\in\mathcal{P}}$ is given by

$$\begin{aligned}\log \mathcal{L}_n\left(b|\Sigma_{\tau_p}\right) = &-\frac{1}{2}\sum_{p\geq 1}\log\left(2\pi k_n^{-1}(\Sigma_{i_p}^{11} + b^2\Sigma_{i_p}^{22} - 2b\Sigma_{i_p}^{12})\right) \\ &- \sum_{p\geq 1}\frac{\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)^2}{2k_n^{-1}\left(\Sigma_{i_p}^{11} + b^2\Sigma_{i_p}^{22} - 2b\Sigma_{i_p}^{12}\right)}.\end{aligned} \tag{88}$$

By (83) note that

$$\log \hat{\mathcal{L}}_n\left(b|\hat{\Sigma}_{i_p}\right) = \log \mathcal{L}_n\left(b|\Sigma_{\tau_p}\right) + o_p(1). \tag{89}$$

Combing these results and the property of stable convergence we see

$$\left(\sqrt{k_n}\epsilon_{i_p}(g,h), \hat{\Sigma}_{i_p}\right)_{p\geq 1} \xrightarrow{\mathcal{L}-s} \left(\varphi_{\tau_p}(g,h), \Sigma_{\tau_p}\right)_{p\geq 1} \tag{90}$$

which implies the convergence of $\log \hat{\mathcal{L}}_n$ to $\log \mathcal{L}$ in probability. This concludes part (a).

(b) We follow the standard procedure of using a mean value decomposition of the log-likelihood. (Note that to ease the notation we drop the dependence of $\hat{\beta}_n^{MLE}(g,h)$ and $\beta_0(g,h)$ on g and h and write simply $\hat{\beta}_n^{MLE}$ and β_0.) Before doing so however we provide functional forms for $\nabla_b \log \hat{\mathcal{L}}_n(b)$ and $\nabla_{bb} \log \hat{\mathcal{L}}_n(b)$ along with associated convergence results. Note that

$$\begin{aligned}\nabla_b \log \hat{\mathcal{L}}_n(b) = &-\frac{1}{2}\sum_{p\geq 1}\frac{2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12}}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}} \\ &+ k_n\frac{1}{2}\sum_{p\geq 1}\frac{\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)^2\left(2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12}\right)}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^2} + k_n\sum_{p\geq 1}\frac{\Delta h(\hat{c}_{i_p})\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}}\end{aligned} \tag{91}$$

and

$$\begin{aligned}\nabla_{bb} \log \hat{\mathcal{L}}_n(b) = &\sum_{p\geq 1} -\frac{2k_n\Delta h(\hat{c}_{i_p})(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p}))(2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12})}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^2} \\ &- \frac{1}{2}k_n(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p}))^2\left(\frac{2(2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12})^2}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^3} - \frac{2\hat{\Sigma}_{i_p}^{22}}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^2}\right) \\ &- \frac{k_n\left(\Delta h(\hat{c}_{i_p})\right)^2}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}} + \frac{1}{2}\left(\frac{(2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12})^2}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^2} - \frac{2\hat{\Sigma}_{i_p}^{22}}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}}\right).\end{aligned} \tag{92}$$

We provide two convergence results that will be used later. To do so let β_n be such that $\beta_n \xrightarrow{\mathbb{P}} \beta_0$. First notice

$$\begin{aligned}
k_n^{-1/2}\nabla_b \log \hat{\mathcal{L}}_n(\beta_0) &= k_n^{-1/2} - \frac{1}{2}\sum_{p\geq 1}\frac{2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12}}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}} \\
&+ \sqrt{k_n}\frac{1}{2}\sum_{p\geq 1}\frac{\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)^2\left(2b\hat{\Sigma}_{i_p}^{22} - 2\hat{\Sigma}_{i_p}^{12}\right)}{\left(\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}\right)^2} + \sqrt{k_n}\sum_{p\geq 1}\frac{\Delta h(\hat{c}_{i_p})\left(\Delta g(\hat{c}_{i_p}) - b\Delta h(\hat{c}_{i_p})\right)}{\hat{\Sigma}_{i_p}^{11} + b^2\hat{\Sigma}_{i_p}^{22} - 2b\hat{\Sigma}_{i_p}^{12}} \\
&\xrightarrow{\mathcal{L}-s} \frac{1}{2}\sum_{p\geq 1}\frac{\varphi_{\tau_p}(g,h)(\Delta g(c_{\tau_p}) - \beta_0\Delta h(c_{\tau_p}))\left(2\beta_0\Sigma_{\tau_p}^{22} - 2\Sigma_{\tau_p}^{12}\right)}{\left(\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}\right)^2} + \sum_{p\geq 1}\frac{\Delta h(c_{\tau_p})\varphi_{\tau_p}(g,h)}{\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}} \\
&\equiv \Xi\left(\beta_0, g, h, (c_{\tau_{p-}}, c_{\tau_{p+}})_{p\geq 1}\right)
\end{aligned} \tag{93}$$

by (79), (82)-(90), and the continuous mapping theorem. Next notice

$$\begin{aligned}
k_n^{-1}\nabla_{bb}\log\hat{\mathcal{L}}_n(\beta_n) &\xrightarrow{\mathbb{P}} \sum_{p\geq 1} -\frac{2\Delta h(c_{\tau_p})(\Delta g(c_{\tau_p}) - \beta_0\Delta h(c_{\tau_p}))(2\beta_0\Sigma_{\tau_p}^{22} - 2\Sigma_{\tau_p}^{12})}{\left(\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}\right)^2} \\
&- \frac{1}{2}(\Delta g(c_{\tau_p}) - \beta_0\Delta h(c_{\tau_p}))^2\left(\frac{2(2\beta_0\Sigma_{\tau_p}^{22} - 2\Sigma_{\tau_p}^{12})^2}{\left(\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}\right)^3} - \frac{2\Sigma_{\tau_p}^{22}}{\left(\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}\right)^2}\right) \\
&- \frac{\left(\Delta h(c_{\tau_p})\right)^2}{\Sigma_{\tau_p}^{11} + \beta_0^2\Sigma_{\tau_p}^{22} - 2\beta_0\Sigma_{\tau_p}^{12}} \equiv H\left(\beta_0, g, h, (c_{\tau_{p-}}, c_{\tau_{p+}})_{p\geq 1}\right)
\end{aligned} \tag{94}$$

by the convergence in probability of β_n to β_0, (79), (82)-(90), and the continuous mapping theorem.

Proceeding with the proof note that

$$\nabla_b \log\hat{\mathcal{L}}_n\left(\hat{\beta}_n^{MLE}|\hat{\Sigma}_{i_p}\right) = 0 \tag{95}$$

since $\hat{\beta}_n^{MLE}$ is the maximizer of $\log\hat{\mathcal{L}}_n$. By the mean value theorem there exists a $\bar{\beta}_n$ falling between $\hat{\beta}_n^{MLE}$ and β_0 such that

$$\begin{aligned}
0 &= \nabla_b \log\hat{\mathcal{L}}_n\left(\hat{\beta}_n^{MLE}\right) \\
&= \nabla_b \log\hat{\mathcal{L}}_n\left(\beta_0\right) + \left(\hat{\beta}_n^{MLE} - \beta_0\right)\nabla_{bb}\log\hat{\mathcal{L}}_n\left(\bar{\beta}_n\right).
\end{aligned} \tag{96}$$

In addition, since $\beta_n^{MLE} \xrightarrow{\mathbb{P}} \beta_0$ and $\bar{\beta}_n$ is bounded between $\hat{\beta}_n^{MLE}$ and β_0, we see $\bar{\beta}_n \xrightarrow{\mathbb{P}} \beta_0$. These results imply

$$\hat{\beta}_n^{MLE} - \beta_0 = -\left(\nabla_{bb}\log\hat{\mathcal{L}}_n\left(\bar{\beta}_n\right)\right)^{-1}\left(\nabla_b\log\hat{\mathcal{L}}_n\left(\beta_0\right)\right) \tag{97}$$

and by (92) and (93)

$$\sqrt{k_n}\left(\hat{\beta}_n^{MLE} - \beta_0\right) \xrightarrow{\mathcal{L}-s} -\left(H\left(\beta_0, g, h, (c_{\tau_{p-}}, c_{\tau_{p+}})_{p\geq 1}\right)\right)^{-1}\Xi\left(\beta_0, g, h, (c_{\tau_{p-}}, c_{\tau_{p+}})_{p\geq 1}\right) \tag{98}$$

concluding the proof. □

Proof of Theorem 3. First, defining $\Omega_n = \{\mathscr{I}_n = \mathscr{I}_n^*\}$, we see $\mathbb{P}(\Omega_n) \to 1$ by Proposition 1. This allows us to restrict the calculations below to the sample paths in Ω_n without loss of generality.

(a) For the ease of exposition we focus on the convergence of $\hat{c}^*_{i_{p+}}$ to $c_{\tau_{p+}}$. The convergence of $\hat{c}^*_{i_{p-}}$ to $c_{\tau_{p-}}$ can be derived analogously and their joint convergence follows readily.

Recall

$$\hat{c}_{i_{p+}} = \frac{1}{k_n \Delta_n} \sum_{m=1}^{k_n} \Delta^n_{i_p+m} X \Delta^n_{i_p+m} X^\top \tag{99}$$

and

$$\hat{c}^*_{i_{p+}} = \frac{1}{k_n \Delta_n} \sum_{m=1}^{k_n} \Delta^n_{i_p+m} X^* \Delta^n_{i_p+m} X^{*\top}. \tag{100}$$

By an application of Lemma B.2 in Dovonon, Gonçalves, and Meddahi (2013) w.p.a.1 we have

$$\left| \frac{1}{k_n \Delta_n} \sum_{m=1}^{k_n} \Delta^n_{i_p+m} X^* \Delta^n_{i_p+m} X^{*\top} - \frac{1}{k_n \Delta_n} \sum_{m=1}^{k_n} \Delta^n_{i_p+m} X \Delta^n_{i_p+m} X^\top \right| \xrightarrow{\mathbb{P}^*} 0 \tag{101}$$

where $\mathbb{P}^*$ is the probability measure induced by the bootstrap sample. Combining these results and the convergence of $\hat{c}_{i_{p+}}$ to $c_{\tau_{p+}}$ under $\mathbb{P}$ we have

$$\begin{aligned} \mathbb{P}\left(\mathbb{P}^*\left(\left|\hat{c}^*_{i_{p+}} - c_{\tau_{p+}}\right|\right)\right) &\leq \mathbb{P}\left(\mathbb{P}^*\left(\left|\hat{c}^*_{i_{p+}} - \hat{c}_{i_{p+}}\right| + \left|\hat{c}_{i_{p+}} - c_{\tau_{p+}}\right|\right)\right) \\ &\to 0 \end{aligned} \tag{102}$$

by (101) and, as $k_n \to \infty$ and $k_n\sqrt{\Delta_n} \to 0$, by Theorem 8.6 in Aït-Sahalia and Jacod (2014).

(b) For each bootstrap sample define the bootstrap log-likelihood as

$$\begin{aligned} \log \hat{\mathcal{L}}^*_n\left(b|\hat{\Sigma}_{i_p}\right) = &-\frac{1}{2}\sum_{p\geq 1} \log\left(2\pi k_n^{-1}(\hat{\Sigma}^{*11}_{i_p} + b^2\hat{\Sigma}^{*22}_{i_p} - 2b\hat{\Sigma}^{*12}_{i_p})\right) \\ &- \sum_{p\geq 1} \frac{\left(\Delta g(\hat{c}^*_{i_p}) - b\Delta h(\hat{c}^*_{i_p})\right)^2}{2k_n^{-1}\left(\hat{\Sigma}^{*11}_{i_p} + b^2\hat{\Sigma}^{*22}_{i_p} - 2b\hat{\Sigma}^{*12}_{i_p}\right)} \end{aligned} \tag{103}$$

where $\Delta g(\hat{c}^*_{i_p})$ and $\Delta h(\hat{c}^*_{i_p})$ are defined as in Algorithm 2 and $\Sigma^*_{i_p} = \Sigma(g, h, \hat{c}^*_{i_{p-}}, \hat{c}^*_{i_{p+}})$ for all $p \geq 1$. By part (a) and the continuous mapping theorem we see

$$\log \hat{\mathcal{L}}^*_n\left(b|\hat{\Sigma}^*_{i_p}\right) = \log \mathcal{L}\left(b|\hat{\Sigma}_{i_p}\right) + o_p(1) \tag{104}$$

since $\log \hat{\mathcal{L}}^*$ and $\Sigma^*_{i_p}$ are continuous in $(\hat{c}^*_{i_{p-}}, \hat{c}^*_{i_{p+}})_{p\geq 1}$. Then, by an analogous argument as in the proof of Theorem 2(a), the result in (104) implies $\hat{\beta}^{*MLE}_n \xrightarrow{\mathbb{P}} \beta_0$.

Since $\hat{\beta}_n^{*MLE}$ is the maximizer of $\log \hat{\mathcal{L}}_n^*$ for each bootstrap sample, by the mean value theorem there exists a $\bar{\beta}_n^*$ falling between $\hat{\beta}_n^{*MLE}$ and $\hat{\beta}_n^{MLE}$ such that

$$\begin{aligned} 0 &= \nabla_b \log \hat{\mathcal{L}}_n^* \left(\hat{\beta}_n^{*MLE} \right) \\ &= \nabla_b \log \hat{\mathcal{L}}_n^* \left(\hat{\beta}_n^{MLE} \right) + \left(\hat{\beta}_n^{*MLE} - \hat{\beta}_n^{MLE} \right) \nabla_{bb} \log \hat{\mathcal{L}}_n^* \left(\bar{\beta}_n^* \right). \end{aligned} \tag{105}$$

Which implies

$$\hat{\beta}_n^{*MLE} - \hat{\beta}_n^{MLE} = - \left(\nabla_{bb} \log \hat{\mathcal{L}}_n^* \left(\bar{\beta}_n^* \right) \right)^{-1} \left(\nabla_b \log \hat{\mathcal{L}}_n^* \left(\hat{\beta}_n^{MLE} \right) \right). \tag{106}$$

Since both $\hat{\beta}_n^{MLE}$ and $\hat{\beta}_n^{*MLE}$ converge in probability to β_0 and $\bar{\beta}_n^*$ falls between $\hat{\beta}_n^{MLE}$ and $\hat{\beta}_n^{*MLE}$ we see $\bar{\beta}_n^* \xrightarrow{\mathbb{P}} \beta_0$ as well. This implies

$$\hat{\beta}_n^{*MLE} - \hat{\beta}_n^{MLE} = - \left(\nabla_{bb} \log \hat{\mathcal{L}}_n^* \left(\beta_0 \right) \right)^{-1} \left(\nabla_b \log \hat{\mathcal{L}}_n^* \left(\beta_0 \right) \right) + o_p(1). \tag{107}$$

Since $\nabla_{bb} \log \hat{\mathcal{L}}_n^*$ and $\nabla_b \log \hat{\mathcal{L}}_n^*$ are continuous in $(\hat{c}_{i_{p-}}^*, \hat{c}_{i_{p+}}^*)_{p \geq 1}$ by the part (a) and the continuous mapping theorem we have

$$\nabla_{bb} \log \hat{\mathcal{L}}_n^* = \nabla_{bb} \log \hat{\mathcal{L}}_n + o_p(1) \text{ and } \nabla_b \log \hat{\mathcal{L}}_n^* = \nabla_b \log \hat{\mathcal{L}}_n + o_p(1). \tag{108}$$

Combining (107) and (108) the result follows by an application of Theorem 2(b).

□

References

Aït-Sahalia, Y., and J. Jacod (2014): *High-Frequency Financial Econometrics*.

Andersen, T. G., O. Bondarenko, and M. T. Gonzalez-Perez (2015): "Exploring Return Dynamics via Corridor Implied Volatility," *Review of Financial Studies*, 28(10), 2902–2945.

Andersen, T. G., N. Fusari, and V. Todorov (2015a): "Parametric Inference and Dynamic State Recovery from Option Panels," *Econometrica*, 83(3), 1081–1145.

Andersen, T. G., N. Fusari, and V. Todorov (2015b): "The pricing of short-term market risk: Evidence from weekly options," *Working Paper*, pp. 1–52.

Andersen, T. G., N. Fusari, and V. Todorov (2015c): "The risk premia embedded in index options," *Journal of Financial Economics*, 117(3), 558–584.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The cross-section of volatility and expected returns," *Journal of Finance*, 61(1), 259–299.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the Macroeconomy, and Asset Prices," *Journal of Finance*, 69(6), 2471–2511.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run," *The Journal of Finance*, LIX(4), 1481–1509.

Black, F. (1976): "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics*, pp. 177–181.

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

Bollerslev, T., N. Sizova, and G. Tauchen (2011): "Volatility in Equilibrium: Asymmetries and Dynamic Dependencies," *Review of Finance*, 16(1), 31–80.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected stock returns and variance risk premia," *Review of Financial Studies*, 22(11), 4463–4492.

Campbell, J., and L. Hentschel (1992): "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, 31(3), 281–318.

Christie, A. A. (1982): "The stochastic behavior of common stock variances. Value, leverage and interest rate effects," *Journal of Financial Economics*, 10(4), 407–432.

Davison, A., and D. Hinkley (1997): *Bootstrap Methods and Their Applications*.

DiCiccio, T. J., and B. Efron (1996): "Bootstrap confidence intervals," *Statistical Science*, 11(3), 189–228.

Dovonon, P., S. Gonçalves, U. Hounyo, and N. Meddahi (2014): "Bootstrapping high-frequency jump tests," *Working Paper*.

Dovonon, P., S. Gonçalves, and N. Meddahi (2013): "Bootstrapping realized multivariate volatility measures," *Journal of Econometrics*, 172(1), 49–65.

Drechsler, I., and A. Yaron (2011): "What Vol Got to Do with It?," *The Review of Financial Studies*, 24(1), 1–45.

Efron, B. (1987): "Better Bootstrap Confidence Intervals," *American Statistical Association*, 82(397), 171–185.

Fama, E., and K. French (1992): "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47(2), 427–465.

Fama, E. F., and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33(1), 3–56.

Figlewski, S., and X. Wang (2000): "Is the 'Leverage Effect' a Leverage Effect?," *Working Paper, Department of Finance, New York University*.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987): "Expected stock returns and volatility," *Journal of Financial Economics*, 19(1), 3–29.

Goncalves, S., and N. Meddahi (2009): "Bootstrapping realized volatility," *Econometrica*, 77(1), 283–306.

Grigelionis, B. (1971): "On the representation of interger-valued random measures by means of stochastic integrals with respect to their Poisson measure," *Lithuanian Mathematics Journal*, 11, 93–108.

Hansen, L. P., and S. F. Richard (1987): "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," *Econometrica*, 55(3), 587–613.

Hou, K., and R. K. Loh (2016): "Have we solved the idiosyncratic volatility puzzle?," *Journal of Financial Economics*, 121(1), 1–28.

Hounyo, U. (2013): "Bootstrapping realized volatility and realized beta under a local Gaussianity assumption," *Working Paper*.

Jacod, J., and P. Protter (2012): *Discretization of Processes*.

Li, J., V. Todorov, and G. Tauchen (2015): "Jump Regressions," *Working Paper*.

Li, J., V. Todorov, G. Tauchen, and R. Chen (2016): "Mixed-scale Jump Regressions with Bootstrap Inference," *Working Paper*.

Mancini, C. (2001): "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, LXIV, 19–47.

Newey, W. K., and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 4, 2111–2245.

Reiss, M., V. Todorov, and G. Tauchen (2015): "Nonparametric test for a constant beta between Ito semi-martingales based on high-frequency data," *Stochastic Processes and their Applications*, 125(8), 2955–2988.

Tauchen, G. (2011): "Stochastic Volatility in General Equilibrium," *Quarterly Journal of Finance*, 01(04), 707–731.

Todorov, V., and G. Tauchen (2011): "Volatility Jumps," *Journal of Business & Economic Statistics*, 29, 356–371.

——— (2012): "Realized Laplace transforms for pure-jump semimartingales," *The Annals of Statistics*, 40(2), 1233–1262.

van der Vaart, A. W. (1998): *Asymptotic Statistics*.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48(4), 817–838.