Variance Risk Premium Dynamics: The Role of Jumps*

Viktor Todorov†

February 6, 2009

Abstract

Using high-frequency stock market data and (synthetic) variance swap rates, this paper identifies and investigates the temporal variation in the market variance risk premium. The variance risk is manifest in two salient features of financial returns: stochastic volatility and jumps. The pricing of these two separate components is analyzed in a general semiparametric framework. The key empirical results imply that investors fears of future jumps are especially sensitive to recent jump activity and that their willingness to pay for protection against jumps increase significantly immediately after the occurrence of jumps. This in turn suggests that time-varying risk aversion, as previously documented in the literature, is primarily driven by large, or extreme, market moves. The dynamics of risk-neutral jump intensity extracted from deep out-of-the money put options confirms these findings. (JEL C51, C52, G12, G13)

Keywords: Continuous-time stochastic volatility models, jump processes, method-of-moments estimation, realized multipower variation.
JEL classification: C51, C52, G12, G13.

---

*This paper is part of my PhD dissertation at the Department of Economics, Duke University. I would like to thank my advisors George Tauchen, Tim Bollerslev, Ron Gallant and Han Hong for many discussions and encouragement along the way. I would like also to thank the editor (Joel Hasbrouck) and anonymous referees for many helpful suggestions. I benefited also from comments, suggestions and discussions with Yacine Ait-Sahalia, Torben Andersen, Ravi Bansal, Alan Bester, Nick Bloom, Mike Chernov, Javier Cicco, Gregory Connor, Darrell Duffie, Rob Engle, Stephen Figlewski, Joel Hasbrouck, Jean Jacod, Ravi Jagannathan, Robert Jarrow, Pedro Duarte, Paul Dudenhefer, Silvana Krasteva, Jonathan Mattingly, Ernesto Mordecki, Mark Podolskij, Barbara Rossi, Albert Shiryaev, Chris Sims, Ken Singleton, Costis Skiadas as well as seminar and conference participants at the Board of Governors, Carnegie-Mellon, Chicago GSB, Duke, Kellogg, LSE, NYU-Stern, Princeton, Stanford GSB, the Conference on Stochastics in Science in Honor of Ole Barndorff-Nielsen, Guanajuato, Mexico, March 2006, the Financial Econometrics conference, Montreal, May 2007.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@kellogg.northwestern.edu

A central topic in finance concerns the premium that investors require for bearing different risks. Much of the work to date has centered on explaining the equity risk premium, i.e. the compensation for the variation in asset prices (price risk). However, price risk is not the only risk investors face when holding assets. Over the last few decades the financial econometrics literature has provided unambiguous evidence that the asset return variances fluctuate over time (see, e.g., the surveys in Bollerslev et al. (1994) and Andersen et al. (2005)). This variation introduces an additional source of risk from holding assets, labeled variance risk. Investors generally dislike the randomness of the future variance and, in equilibrium, demand a premium for accepting this risk. This gives rise to the so-called variance risk premium.

The presence of the variance risk premium at the aggregate market level has already been extensively documented in the literature.[1] Bakshi and Kapadia (2003) detect in a nonparametric way the presence of variance risk premium by analyzing the profits and losses from delta-hedged positions in S&P 500 and S&P 100 index options. Carr and Wu (2009) calculate a model-free measure of *realized* variance risk premia by comparing the realized variance with its risk-neutral expectation. They show that this measure is on average negative for a range of stock market indexes, which indicates that market variance risk is indeed priced. Bakshi and Madan (2006) link the variance risk premium with higher order moments of the return distribution and investor risk aversion. Additional evidence for presence of a variance risk premium are also available from fully parametric estimation of the pricing kernel, see, e.g., Bates (2000), Chernov and Ghysels (2000), Pan (2002) and Eraker (2004).

Although the existence of a (market) variance risk premium is by now well established, much less is known about the dynamic dependencies in the premium. This is the main focus of the present paper. In particular, does the compensation demanded by investors for market variance risk change over time? If so, what determines this variation? Is it possible to associate the temporal variation with observable characteristics of the market, including past variances and/or jumps? Is it possible to reconcile the evidence for a time-varying variance risk premium with a valid economy-wide pricing kernel, and in turn what are the properties of such a pricing kernel? These are the questions the current study seeks to answer.

The paper starts by identifying the sources of market variance risk. The reasons for the *realized* market variance to change over time are twofold. First is the well-known presence of stochastic volatility, i.e. *ex-ante* (expected) volatility changes over time. Second, variance risk is also associated with the occurrence of *unanticipated* market jumps. In contrast to previous work, which has typically associated variance risk with the presence of stochastic volatility only, this study adopts a general semiparametric stochastic volatility model that explicitly allows for both sources of variance risk. Since there is less agreement on the correct *parametric* model for the jumps in the price and the stochastic volatility as well as their dependence, I purposely leave this part of the model unspecified. On the other hand, previous empirical work has convincingly demonstrated the need for a two-factor structure akin to the one adopted

[1] Risk premium for individual stocks variance has been analyzed in Bakshi et al. (2003) and Carr and Wu (2009).

here for satisfactory describing the dynamic dependencies in the stochastic volatility, e.g., Andersen et al. (2002), Alizadeh et al. (2002), Chernov et al. (2003). This semiparametric modeling approach thus strikes a reasonable balance between precision of the estimates for the variance risk premium on the one hand and robustness of the results on the other.

To actually estimate the variance risk premium, I further explore the recent advances in high-frequency financial data analysis. In particular, using 5-minute high-frequency data on the S&P 500 index, I construct *model-free* measures of the realized variance and the realized jumps for each of the days in the sample. To estimate the parameters of the stochastic volatility model, I then match moments of these statistics with the ones implied by the model. The result of the estimation is a model-based measure for the future expected variance under the true probability. The variance risk premium estimate is then simply constructed by differencing this measure with the model-free risk-neutral expectation of the future variance. The latter is inferred from a portfolio of out-of-the money option prices on the S&P 500 index that synthesizes variance swap.[2]

The resulting estimated variance risk premium varies quite considerably over time. The estimation results also reveal that both past "realized" price jumps and stochastic volatility are important determinants of the variance risk premium dynamics. Importantly, the estimated variance risk premium does not rely on any specific pricing kernel, only requiring the model-free risk-neutral expectation of the future variance.

To further check if the inferred variance risk premium dynamics can be rationalized in a no-arbitrage setting, I model theoretically the compensation demanded by the investors for each of the sources of variance risk. Keeping with the semiparametric approach adopted in the paper, I specify only the part of the economy-wide pricing kernel that determines the prices of these risks. The modeling of the jump risk price in particular is quite flexible as it allows for the jumps to be time-homogenous under the physical measure, yet exhibit significant persistence under the risk-neutral measure. I infer the dynamics of the pricing kernel, by matching moments of the variance risk premium against those implied by the pricing kernel.

The results suggest that jumps play a very important role in explaining the variance risk premium. Jumps are clearly present in the level of market prices, and when a jump occurs, this is typically associated with a spike, or a jump, in the stochastic volatility as well. The effect of jumps on the future market dynamics is limited, however. The jump volatility factor has a very quick mean reversion, thus capturing the higher frequency moves in the volatility, while the very persistent changes in volatility are driven by the continuously evolving volatility factor.

Meanwhile, the estimated variance risk premium typically increases after a big market jump and slowly reverts to its long-run mean thereafter. This is explained with compensation for jump risk that depends on a persistent state variable which, in turn, is related with the "realized" price jumps. The differential impact of jumps on the risk premium and the dynamics of the underlying asset therefore suggests that investor attitude towards jumps is time-varying. Recent empirical evidence in

[2] This latter measure also coincides with the new VIX index computed from CBOE.

favor of time-varying risk aversion have also been reported by, e.g., Bollerslev et al. (2005) and Brandt and Wang (2003). In contrast to these earlier studies, the results reported here explicitly link the changes in risk aversion with extreme market events, or crashes, and suggest that it is the attitude towards these big changes that gets revised immediately after such events occur.

To further corroborate the main empirical findings, I conduct two robustness checks. The first relates to presence and structure of jumps in the returns and stochastic volatility. The stochastic volatility is not directly observable. Instead, it is possible to use the *observable* VIX index. The latter represents a risk-neutral expectation of the future variance. Given the fact that both the variance risk premium and the future variance depend on the current stochastic volatility, a jump in the latter will translate directly into a jump in the VIX index. Using the test proposed by Lee and Mykland (2008), I verify the existence of a nontrivial number of days in the sample for which there is a jump in the stochastic volatility. Further, on a significant number of these days the market also appears to jump.

The second robustness check concerns the impact of past "realized" jumps on the future compensation for jump risk demanded by investors. This is the channel through which past jumps impact the variance risk premium. To isolate the compensation of jump risk, I use deep out-of-the-money close-to-maturity options on the S&P 500 index and extract a model-free measure for the future tail jump intensity under the risk-neutral distribution. This measure shows strong dependence on the past realized jumps, directly in line with the predictions of the model implied by the pricing kernel.

The remainder of the paper is organized as follows. Section 1 introduces the general stochastic volatility model for the underlying asset under the physical measure, along with the actual fit to the S&P 500 data. Section 2 constructs the variance risk premium and reports empirical evidence for temporal variation in this measure. Section 3 derives general prices of diffusive and jump risk within the stochastic volatility model and discusses their implication for the variance risk premium. This section also presents tests for the different specifications of diffusive and jump risk compensation using high-frequency data on the underlying asset and the VIX index. Section 4 details the nonparametric robustness checks. Section 5 concludes. The main proofs are given in an appendix at the end of the paper, with additional details available in a supplementary appendix.

1 Dynamics under the Physical Measure

I start the analysis of the (market) variance risk premium by identifying the sources of variance risk in the context of a general semiparametric stochastic volatility model. In this section I also estimate the model using only high-frequency data on the underlying index in order to: (1) check if it has problems fitting this data alone and (2) construct later an estimate for the conditional expected future variance to be compared with the one under the risk-neutral measure (the VIX index).

1.1 The Stochastic Volatility Model

I fix a filtered probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with $\mathbf{F} = (\mathscr{F}_t)_{t\in\mathbb{R}}$ denoting the filtration. On this space I define with $F(t)$ the price at time t of a futures contract on the stock market index expiring at some date in the future. I assume that $f(t) := \log(F(t))$ has the following dynamics under the physical measure $\mathbb{P}$

$$df(t) = b(t)dt + \sigma(t)(\rho dB_1(t) + \sqrt{1-\rho^2}dB_2(t)) + \int_{\mathbb{R}_0^n} h(\mathbf{x})\tilde{\mu}(dt, d\mathbf{x}), \tag{1}$$

$$\sigma^2(t) = V^c(t) + V^j(t), \tag{2}$$

$$dV^c(t) = \kappa_c(\overline{V}^c - V^c(t))dt + \sigma_{cv}\sqrt{V^c(t)}dB_1(t), \tag{3}$$

$$dV^j(t) = \rho_j V^j(t)dt + \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(dt, d\mathbf{x}), \tag{4}$$

where (B_1, B_2) is a standard Brownian motion; $\mathbf{x}$ is an n-dimensional vector on $\mathbb{R}_0^n$; μ is a time-homogenous Poisson random measure with compensator (intensity) ν such that $\nu(dt, d\mathbf{x}) = dtG(d\mathbf{x})$ for some $G : \mathbb{R}_0^n \to \mathbb{R}_+$; $h : \mathbb{R}_0^n \to \mathbb{R}$ and $k : \mathbb{R}_0^n \to \mathbb{R}_+$ and $\tilde{\mu} := \mu - \nu$ is the compensated measure.

The futures price in (1) has three components. The first is the drift term, $b(t)$, which is left unspecified in this paper. The second component of the price is a continuous martingale. Its time-variation is determined by the process $\sigma^2(t)$. I refer to $\sigma^2(t)$ as the stochastic variance and model it as a sum of two factors. The first one, $V^c(t)$, is the continuous component of the stochastic variance and follows a square-root process as in the standard affine stochastic volatility models (Duffie et al. (2000)). The second component of the stochastic variance, $V^j(t)$, is its discontinuous part. $V^j(t)$ is a non-Gaussian OU process (Barndorff-Nielsen and Shephard (2001)). $V^j(t)$ can be equivalently written as

$$V^j(t) = \sum_{s\leq t} e^{\rho_j(t-s)}\Delta V^j(s),$$

and therefore it is simply a moving-average of past variance jumps. The impact of the past jumps on the current level of $V^j(t)$ (and hence $\sigma^2(t)$) is determined by the value of ρ_j. The last component of the price in equation (1) is a jump martingale. Intuitively this component of the price is just a compensated (i.e. demeaned) sum of jumps.

The variance of $f(t)$ is measured by its quadratic variation (hereafter QV). For a period $(t, t+a]$ it is given by

$$[f, f]_{(t,t+a]} = \int_t^{t+a} \sigma^2(s)ds + \sum_{t<s\leq t+a} (\Delta f_s)^2. \tag{5}$$

The first term in (5) is due to the continuous part of the futures price. I refer to it as Integrated Variance (hereafter IV) and denote it $IV_a(t) := \int_t^{t+a} \sigma^2(s)ds$. The second term in (5) is the quadratic variation of the jumps.

The randomness (or time-variation) in QV generates *variance risk* which we are after. It is present either because the conditional expected future volatility is random (i.e. $\sigma^2(t)$ depends on t) or because there are jumps in the price. Note that the second term in (5) is random in spite of the fact that the price jumps are time-homogenous (i.e. their conditional volatility is constant). This is very different from the behavior of the quadratic variation of the continuous martingale, which in the case of time-homogeneity (i.e. $\sigma^2(t) \equiv const$) is non-random.

I finish this section with a short discussion of the empirically relevant features of the SV model (1)-(4), which also explains why this model is chosen.

Price Jumps. The inclusion of price jumps is a necessity rather than a modeling choice. Indeed, recently Barndorff-Nielsen and Shephard (2006), Jiang and Oomen (2008), Ait-Sahalia and Jacod (2009), Lee and Mykland (2008) and Jacod and Todorov (2009) propose different nonparametric tests for presence of price jumps based on high-frequency data. All these papers find overwhelming evidence for jumps in asset prices which is in line with earlier parametric evidence. In the model (1)-(4), the price jumps are time-homogenous, i.e., they are modelled as a Lévy process. In the empirical part in Section 1.2 I find that for the data used in this paper there is no significant time-variation in the price jumps and hence their modeling with a Lévy process.

Stochastic Volatility. Another important feature of the financial data is the persistence in the returns variance. Since in the model here the price jumps are time-homogenous, their conditional variance is constant. Therefore, persistence in the returns variance can be generated only through time-variation in $\sigma^2(t)$. In the model here $\sigma^2(t)$ is a sum of two factors, each of which is an AR(1)-type process. This provides a parsimonious way of capturing persistence in the volatility as already shown in previous empirical work and it will be further verified in Section 1.2. An important feature of the model (1)-(4) is that the stochastic variance contains jumps. Later in Section 4, I will provide further nonparametric evidence for their presence.

Jump Dependence. The modeling of the jumps in the price and the variance is quite flexible. In equations (1) and (3) the jump sizes in the price and the variance are expressed as functions of jumps in an n-dimensional space. This way all possible dependencies between the jumps can be generated. Examples include independent price and variance jumps, perfect linear dependence between price and variance jumps (Barndorff-Nielsen and Shephard (2001)) and variance jumps being proportional to the squared price jumps (Todorov (2009a)).

For the purposes of the analysis in this paper I leave $h(\cdot)$, $k(\cdot)$ and $G(\cdot)$ unspecified. While the dependence between price and variance jumps is hard to detect with low-frequency data (see e.g. Eraker (2004), Broadie et al. (2007)), this is not the case when high-frequency data is used (as illustrated below). Therefore, since we do not have a clear idea of the dependence structure of the jumps, it is better to leave this part of the model unspecified and let the data "choose" the right one. This way potential misspecification problems can be avoided.

Leverage Effect. Although this study is not concerned with the "leverage effect" (i.e. the (negative) linear relationship between the price and variance innovations), in order for the model to be empirically realistic it should allow for a flexible way of generating it and this is the case here. Since the price and the variance are both driven by Brownian motions as well as jumps, the "leverage effect" could be generated in two different ways in this model. One way is through correlating the Brownian motions and the second way is through linking the jumps in the price and the variance.

1.2 Estimation of the Model under the Physical Measure

I proceed with estimating the model (1)-(4). The estimation is done using high-frequency data and is based on the following idea. First, I aggregate the high-frequency data into daily measures of QV and IV. These measures are *model-free* and asymptotically (as we sample more and more frequently) they converge to their unobservable counterparts. Then, I estimate the model using GMM by simply matching sample moments of QV and IV with those implied by the model. Apart from being easy to implement, the outlined estimation strategy provides a robust way to identify the parameters controlling the jumps and the stochastic volatility since the inference is based directly on QV and QV-IV, which separate stochastic volatility from jumps. Thus, we do not leave deeply structural parameters to do this separation, but rely solely on the data for this. In addition, the model-free measures of QV and IV will provide us later in the analysis with an easy and robust way to isolate the effect of jumps on the temporal variation in the risk premium.

The details are as follows. The unit of measurement in this paper is one business day. The business day starts from the close of trading of the previous day. It consists of two parts: the first part is the overnight, which is the time till the opening of the trading. The second part of the business day is the trading period. The first part of the business day is a fraction of π and the second is $1-\pi$. Using high-frequency observations of the price process during the trading part of the day and applying results in Barndorff-Nielsen and Shephard (2006) and Barndorff-Nielsen et al. (2005), I construct model-free measures for $IV(t) := IV_{1-\pi}(t+\pi)$ and $QV(t) := QV_{1-\pi}(t+\pi)$.[3] For QV I use the so-called Realized Variance (hereafter abbreviated as RV) defined over the trading part of day t as

$$RV_\delta(t) = \sum_{i=1}^{M} r_\delta^2(t+\pi+(i-1)\delta), \tag{6}$$

where δ is the length of the high-frequency interval (e.g. 5-min.), $M = \lfloor 1/\delta \rfloor$ is the number of intra-day observations and $r_\delta(t) := f(t+\delta) - f(t)$ for any t. The estimate for IV is the so-called Realized Tripower Variation (hereafter abbreviated as TV), which is defined over the trading part of day t as

$$TV_\delta(t) = 1.9358 \sum_{i=3}^{M} |r_\delta(t+\pi+(i-3)\delta)|^{2/3} |r_\delta(t+\pi+(i-2)\delta|^{2/3} |r_\delta(t+\pi+(i-1)\delta)|^{2/3}. \tag{7}$$

Intuitively, the use of consecutive returns in TV "kills" the jumps and therefore TV measures only the continuous part of QV, i.e. IV. From now on I treat $TV_\delta(t)$

[3] $IV(t)$ and $QV(t)$ are IV and QV respectively over the trading part of the business day t.

and $RV_\delta(t)$ as their unobservable counterparts $IV(t)$ and $QV(t)$ respectively. Under certain regularity conditions, as shown in Todorov (2009b), this has no asymptotic effect on the estimation results in this paper.

Since our interest here is only in the behavior of QV, I estimate only parameters controlling it. To avoid identification problems in the estimation, the model is re-parameterized as follows. I set $\eta := \overline{V}^c - \frac{1}{\rho_j}\int_{\mathbb{R}_0^2} k(\mathbf{x})G(d\mathbf{x})$ and $\sigma_c := \sigma_{cv}\sqrt{\frac{\overline{V}^c}{2\kappa_c}}$. η is the mean of the stochastic variance and σ_c^2 is the variance of the diffusive variance component $V^c(t)$. In the estimated model I do not parameterize the distribution of the jumps in the price and the variance. Instead, I treat as parameters only cumulants which are needed for computing the moment conditions in the GMM. In particular, I estimate the following cumulants[4]

$$\int_{\mathbb{R}_0^n} h^2(\mathbf{x})G(d\mathbf{x}), \quad \int_{\mathbb{R}_0^n} h^4(\mathbf{x})G(d\mathbf{x}), \quad \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(d\mathbf{x}) \quad \text{and} \quad \int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}).$$

The advantage of estimating only cumulants of the jumps is robustness. The parameters controlling the jumps are the hardest to estimate and there is no consensus in the literature about the correct parametric model and even less so about the jump dependence. Therefore, a wrong parametric specification or unidentified deeply structural jump parameters can affect the correctness of all estimated moments of QV. This in turn will lead to a wrong conditional expectation of future QV and its dependence on the past information, which we are after. Estimating the cumulants of the jumps avoids all these problems, as it reflects directly what is in the data.

Turning to the moment conditions in the GMM, I match the following statistics: the ratio of the mean of QV to the mean of squared daily returns; mean, variance and autocorrelation of IV; mean and variance of QV; mean of Realized Fourth Variation (hereafter FV), which is defined in (8) below. For the autocorrelation of IV I use lags 1, 3 and 6 as well as the average autocorrelation for lags $11-20$, $21-30$ and $31-40$. The averaging of the higher order autocorrelations is done since these autocorrelations are estimated with less precision. Altogether I end up with 11 moment conditions. The Realized Forth Power Variation is defined for a day t as

$$FV_\delta(t) = \sum_{i=1}^{M} r_\delta^4(t + \pi + (i-1)\delta). \tag{8}$$

FV is a measure of the sum of the price jumps raised to the power four over the day and its mean helps identify the second order moments of the jumps in the price and the variance. The moments used in the estimation are calculated using results

[4] In the estimation of the model I impose the following constraint on the cumulants

$$0 \le \int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}) \le \sqrt{\int_{\mathbb{R}_0^n} h^4(\mathbf{x})G(d\mathbf{x}) \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(d\mathbf{x})},$$

which guarantees the existence of a two-dimensional Lévy process (for the jumps in the price and the variance) with cumulants equal to the estimated ones.

in Todorov (2009a). Further details on the estimation are provided in a separate appendix.

I conclude this section with some details on the high-frequency data used in the estimation. The data is on the S&P 500 index futures contract for the period January 2, 1990, to November 29, 2002. There are 80 five-minute return observations in each day covering the day trading session from 9:30am till 4:15pm. For each of the days in the sample I calculate RV and TV and their difference JV:=RV-TV, which is a measure of squared jumps over the day. Figure 1 plots the daily returns as well as TV and JV series. As seen from the TV series, integrated variance has spikes and this is suggestive of jumps in the stochastic variance $\sigma^2(t)$. Another interesting observation from Figure 1 is that most of the days in which TV is high are days in which JV is high as well. Finally, the two series differ significantly in their persistence, as seen from their first 100 autocorrelations plotted on Figure 2: IV is a very persistent process unlike the squared price jumps.

1.3 Estimation Results

The estimation results are reported in Table 1. In what follows I summarize the key findings from the estimation. First, the test of overidentifying restrictions shows that the model provides good fit to the high-frequency data. This is further confirmed by the diagnostic t-statistics associated with each of the moments used in the estimation reported in Panel B of Table 1. Under correct model specification these statistics should be approximately standard normal and thus large values signal difficulty in matching the particular moment condition. The results in Table 1 suggest that the model has no problem with fitting any particular moment used in the estimation.

The two-factor structure of the stochastic variance $\sigma^2(t)$ yields empirically plausible persistence in IV. Figure 3 plots the fit to the autocorrelation of TV, implied by the parameter estimates. As seen from the figure, the autocorrelation of TV is well matched for lags until forty. After lag forty the model-implied autocorrelation slightly underestimates the empirically observed one. However, it is still well within the 95% confidence interval. On the other hand more parsimonious one-factor restrictions of the current model cannot match the autocorrelation in IV. This holds regardless of which of the two variance factors is excluded. First I test that the continuous variance factor is not present, i.e. that $V^c(t) \equiv 0$ which is equivalent to test $\sigma_c \equiv 0$. Note that under the null hypothesis κ_c is not identified. I use a criterion difference test, i.e $J(\widehat{\theta}_r) - J(\widehat{\theta})$, where $J(\theta) = T\widehat{m}_T(\theta)'\widehat{W}m_T(\theta)$ is the GMM objective function and $\widehat{\theta}_r$ and $\widehat{\theta}$ are the restricted and unrestricted estimates respectively. From the results in Andrews (2001), it follows that the criteria different test has a χ^2_1 limiting distribution under the null.[5] The value of the test is 96.1582. This shows

[5]Assumption 3 in Andrews (2001) is easy to be verified in our setting. This assumption is needed for proving the asymptotic distribution of the criterion difference test. It concerns the convergence of the first derivative of the GMM objective function, as a function of the nuisance parameter present only under the alternative κ_c (the convergence is on the space of continuous functions of κ_c equipped with the uniform metric). The first derivative of the vector of moment conditions is a continuous function of κ_c and the moment vector does not depend on κ_c under the null hypothesis. Functional

that such restriction is implausible. Similarly testing $V^j(t) \equiv 0$ is equivalent to testing $\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(d\mathbf{x}) \equiv \int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}) \equiv 0$. As in the previous test we have a nuisance parameter identified only under the alternative. In this case this is ρ_j. I use again a criterion difference test which is now asymptotically χ^2_2. The value of the test is 140.8002. Note the huge value of the test. By excluding the discontinuous component we miss not only the autocorrelation in IV, but we also cannot generate enough volatility of the stochastic volatility to match the observed data. This deficiency of the square-root process is also documented in Eraker et al. (2003) who use lower frequency stock market data and in Broadie et al. (2007) who use options data. Thus for the purposes of constructing a correct estimate for the variance risk premium and identifying its dynamic dependencies we need the two-factor structure of $\sigma^2(t)$.

The estimated cumulants of the Lévy process show that there is a strong relationship between the jumps in the variance and the jumps in the price, i.e. $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x})$ is statistically different from zero. Importantly, this finding rejects independence between price and variance jumps and this fact has important implications for identifying and interpreting the dynamic dependence of the variance risk premium on the variance factors and price jumps as will become clear later in Sections 2 and 3. This result is also in line with the observation made at the beginning of the current section regarding the positive link between JV and TV series. It should be noted that most popular ways of modeling the dependence between the price and variance jumps that have been estimated before, see e.g. Eraker et al. (2003), can be rejected. As already discussed in Section 1 here I do not model parametrically the link between the jumps in the variance and the jumps in the price. Therefore, the results in the paper are not driven by a (possibly misspecified) parametric model for the jumps but rather reflect what is in the data.

Overall, the tests conducted above show that the model (1)-(4) can successfully match the observed high-frequency data, while any restricted version of it (e.g. by removing factors of $\sigma^2(t)$ or forcing independence between variance and price jumps) will miss important features of the stock market index data. In line with many other studies in the literature (Andersen et al. (2002), Alizadeh et al. (2002), Chernov et al. (2003)) here I find one of the variance factors to be slowly mean reverting, having a half-life of approximately twenty (business) days, while the other one to be quickly mean reverting with a half-life of approximately half a day. Perhaps not surprisingly the quickly mean-reverting factor is the jump component of the variance, while the slowly mean-reverting variance factor is the continuous component of the variance.

What is the economic interpretation of these results? The estimation shows that the market index possesses strong persistence in the variance and occasional jumps. The occurrence of price jumps has no effect on the future price level, but the estimation results show that it leads to an increase of the stochastic volatility. This

convergence of the first derivative of the GMM objective function, as a function of κ_c then follows by establishing finite-dimensional convergence (which in turn follows essentially from a standard CLT theorem for the moment vector) and C-tightness of the sequence, which can be verified using Theorem 8.3 in Billingsley (1968).

effect of the price jumps on the volatility is relatively short-lived and is on the top of the slowly changing level of volatility, which is captured by the continuous volatility factor.

2 Initial Analysis of the Variance Risk Premium

Given the good fit of the SV model (1)-(4) to the high-frequency data, we are now ready to start the analysis of the variance risk premium. The variance risk premium is the compensation for variance risk (which is measured by the quadratic variation) and is therefore defined as the wedge between (conditional) expectation of the future quadratic variation (of the index) under the physical and the risk-neutral measure. Thus, the daily-standardized risk premium for the variance risk over the next a days is

$$VR_a(t) = \frac{1}{a}\mathbb{E}^{\mathbb{P}}\left([f,f]_{(t,t+a]}|\mathscr{F}_t\right) - \frac{1}{a}\mathbb{E}^{\mathbb{Q}}\left([f,f]_{(t,t+a]}|\mathscr{F}_t\right), \tag{9}$$

where $\mathbb{E}^{\mathbb{Q}}(\cdot)$ denotes expectation under the risk-neutral measure.[6]

In this Section I construct a measure for the variance risk premium, using VIX index data and the SV model (1)-(4), and analyze the dynamics of the variance risk premium.

2.1 Measuring Variance Risk Premium

I start with constructing a measure for the variance risk premium. We can recognize the first term in (9) as the (daily-standardized) variance swap rate on the S&P 500 index (i.e. the price of a forward contract on the future variance) which we denote as $SW_a(t) := \frac{1}{a}\mathbb{E}^{\mathbb{Q}}\left([f,f]_{(t,t+a]}|\mathscr{F}_t\right)$. Following Bakshi and Madan (2000), Britten-Jones and Neuberger (2000), Carr and Wu (2009), its value can be inferred (synthesized) from the following (static) portfolio of options

$$\frac{2}{a}\sum_i e^{ar_{(t,t+a]}}\frac{\Delta K_i}{K_i^2}O(K_i,a),$$

where $O(K_i,a)$ is an out-of-the money put or call option price on the S&P 500 contract (of European style) with time to expiration of a days; $r_{(t,t+a]}$ is the interest rate over the period $(t,t+a]$; $\Delta K_i = K_i - K_{i-1}$ and the summation in the above formula is over all available strike levels.[7] The above portfolio of options has been used in the

[6] In case a superscript is not put on the expectation operator, the expectation is always assumed to be under the physical measure.

[7] As shown in Carr and Wu (2009), when the price contains jumps there is an error in the replication (in addition to the discretization one). For the model (1)-(4), this error is $\epsilon_{t,a} = \frac{-2}{a}\int_t^{t+a}\int_{\mathbb{R}_0^n}\left(e^{h(\mathbf{x})} - 1 - h(\mathbf{x}) - \frac{h^2(\mathbf{x})}{2}\right)\nu^{\mathbb{Q}}(ds,d\mathbf{x})$, where $\nu^{\mathbb{Q}}(\cdot,\cdot)$ is the compensator of the jump measure μ under the risk-neutral measure. The numerical experiments in Carr and Wu (2009) suggest that the error is not significant for practical purposes. In fact, since in this paper I am interested only in the time-variation of the variance risk premium, taking into account this error does not change any of the conclusions.

calculation of the new VIX index (with a corresponding to one month), which we can therefore view as a portfolio of short-maturity options. Thus, I use the VIX index to compute the first term in (9).[8]

To compute the variance risk premium we are left with estimating the conditional expectation under $\mathbb{P}$ of the future QV. I use the model (1)-(4) for this. For the discontinuous component of the quadratic variation this is easy. Its conditional expectation is equal to the unconditional one since the price jumps are time-homogenous. Turning to the continuous part of the quadratic variation, its conditional expectation is different from the unconditional one since $\sigma^2(t)$ is time-varying. The conditional expectation of IV is a linear function of its two variance factors. However, this conditional expectation is not available to the econometrician. At the same time, the model implies an ARMA(2,2) process for the daily IV with coefficients determined from the structural parameters of the model. This fact could be used to calculate the linear projection of the integrated variance on the past values of TV and JV (which in turn are being used as a proxy for IV and QV-IV respectively). The details are provided in a separate appendix to this paper. I denote this linear projection with $P(IV_a(t)|\mathscr{G}_t)$, where $\mathscr{G}_t = \sigma(TV_\delta(t-1), ..., TV_\delta(1), JV_\delta(t-1), ..., JV_\delta(1))$, i.e. $\mathscr{G}_t$ is the information created from the past realizations of TV and JV, and we have $\mathscr{G} \subset \mathscr{F}$.

Thus, a feasible measure for the premium at date t for the variance risk over the next a days is

$$RP_a(t) = \int_{\mathbb{R}_0^n} h^2(\mathbf{x})G(d\mathbf{x}) + \frac{1}{a}P(IV_a(t)|\mathscr{G}_t) - SW_a(t). \tag{10}$$

Its usefulness comes form the following[9]

$$\mathrm{Cov}(RP_a(t), TV_\delta(t-j)) = \mathrm{Cov}\left(VR_a(t), IV(t-j)\right), \tag{11}$$

$$\mathrm{Cov}(RP_a(t), JV_\delta(t-j)) = \mathrm{Cov}\left(VR_a(t), QV(t-j) - IV(t-j)\right). \tag{12}$$

Below I illustrate how we can make use of these two sets of covariances to identify the dynamics of the variance risk premium. The natural starting point is the case of a constant variance risk premium. In this case we simply have

$$\begin{aligned} SW_a^c(t) :&= \frac{1}{a}\mathbb{E}\left(IV_a(t) + \int_t^{t+a}\int_{\mathbb{R}_0^n} h^2(\mathbf{x})G(d\mathbf{x}) \middle| \mathscr{F}_t\right) + K, \\ &= K_0 + \frac{1-e^{-\kappa_c a}}{a\kappa_c}V^c(t) + \frac{e^{\rho_j a}-1}{a\rho_j}V^j(t), \end{aligned} \tag{13}$$

[8]In the calculation of the VIX index a calendar-counting convention is used. That is, the year consists of 365 days and in computing the time to expiration for the options, the actual number of days is being used. However, in this paper I adopted a business-time counting. That is, a unit of time here is one business day. I continue to use the business-time convention and assume that each month consists of 22 business days. I use the VIX index to calculate a daily-standardized variance swap rate with one month horizon (corresponding to 22 business days): $SW_{22}(t) = \frac{30}{365}\frac{1}{22}VIX^2(t)$.

[9](11)-(12) hold if an asymptotic approximation for TV and JV given in the separate appendix to this paper holds exact; otherwise they hold asymptotically as $\delta \to 0$.

where K is the constant variance risk premium and K_0 is some constant. This means that in the case of a constant variance risk premium the variance swap rate is a linear combination of the variance risk factors. Importantly, the coefficients in front of the variance factors in $SW_a^c(t)$ are not free and are determined by the persistence of these factors *under the physical measure*. A natural generalization is to consider variance risk premium specification under which the variance swap rate is a linear combination of the variance factors, but the coefficients in front of them are not restricted. This corresponds to a variance risk premium which is linear in the variance factors and the next section provides prices of diffusive and jump risk that support such variance risk premium specification. Thus, under this generalization of the variance risk premium, the variance swap rate is

$$SW_a^v(t) := K_0 + K_c V^c(t) + K_j V^j(t), \tag{14}$$

where K_0, K_c and K_j are some constants. The constant K_0 in (13) and (14) might differ of course. All parametrizations of the pricing kernel considered previously in the literature would imply (14) for the variance swap rate (with some further restrictions on the coefficients in most cases). The variance swap rate corresponding to constant variance risk premium can be recovered by constraining $K_c \equiv \frac{1-e^{-\kappa_c a}}{a\kappa_c}$ and $K_j \equiv \frac{e^{\rho_j a}-1}{a\rho_j}$ in equation (14).

Can we distinguish these two scenarios for the variance risk premium using $RP_a(t)$? The answer to this question is positive and the reason for this is that we have access to IV and QV (measured from the high-frequency data by TV and RV respectively). To show this I work with the variance swap specification $SW_a^v(t)$ in (14), since $SW_a^c(t)$ is a constrained version of it. Then, it is easy to prove that for $i = 1, 2, ...$

$$\mathrm{Cov}(VR_a(t), IV(t-i)) = 0 \quad \Leftrightarrow \quad K_c \equiv \frac{1 - e^{-\kappa_c a}}{a\kappa_c} \quad \text{and} \quad K_j \equiv \frac{e^{\rho_j a} - 1}{a\rho_j}.$$

In other words, provided there is time-varying risk premium with time-variation determined by the variance factors, $VR_a(t)$ should be correlated with the past values of IV. Further, we can investigate whether both the jump and diffusion parts of the stochastic variance $\sigma^2(t)$ determine the time-variation in the variance risk premium. It is easy to derive that for $i = 1, 2, ...$

$$\mathrm{Cov}(VR_a(t), QV(t-i) - IV(t-i)) = 0 \quad \Leftrightarrow \quad K_j \equiv \frac{e^{\rho_j a} - 1}{a\rho_j} \quad \text{or} \quad \int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}) \equiv 0.$$

Since the empirical results in Section 1.2 indicate that $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}) \neq 0$, $VR_a(t)$ will be correlated with the past squared price jumps provided the variance risk premium depends on the variance jump factor $V^j(t)$. Thus, with the covariance between RP and TV and RP and JV, we can differentiate constant variance risk premium from variance risk premium that is linear in the variance factors. Further, because of the link between the price and variance jumps, using these covariances, we can also determine which of the variance factors determines the variation in the variance risk premium.

2.2 Dynamics of the Variance Risk Premium

From (9) we can easily get

$$\mathbb{E}(VR_a(t)) = \frac{1}{a}\mathbb{E}([f, f]_{(t,t+a]}) - \mathbb{E}(SW_a(t)),$$

$$\sqrt{\mathrm{Var}\,(VR_a(t))} \geq \left| \frac{1}{a}\sqrt{\mathrm{Var}\,\left(\mathbb{E}([f, f]_{(t,t+a]}|\mathscr{F}_t)\right)} - \sqrt{\mathrm{Var}(SW_a(t))} \right|.$$

Thus, using the parameter estimates of the SV model as well as the variance swap rate data, we have an estimate for the first two moments of the variance risk premium. Its estimated mean is -0.4015, while the mean of the variance swap rate is 1.6542 (both in daily variance units). This shows that the variance risk premium is rather significant, which is consistent with earlier studies of Bakshi and Kapadia (2003) and Carr and Wu (2009). An estimate for the lower bound of the volatility of the variance risk premium is 0.4150 and this is to be compared with the estimated volatility of the variance swap which is 1.1303.

The above evidence suggests that the variance risk premium is not only big but it also shows significant variation over time. This has important consequences. First, this implies that the wedge between the physical and risk-neutral measure is pretty wide. Secondly, this wedge is not constant. Third, the variance risk factors have more variance under the risk-neutral measure, and this suggests that changes in investors expectations about the significance of these risk factors account for this wedge. Finally, from a practical point of view, the nontrivial time-variation in the variance risk premium means that for the purposes of pricing derivatives on the volatility, like variance swaps, one can not use directly the voluminous literature on modeling and forecasting volatility (which is under the $\mathbb{P}$ measure). One also needs to know what drives the variance risk premium.

I finish this section with an analysis of the measure RP (evaluated at the parameter estimates reported in Table 1). Figure 4 plots the RP series together with TV and JV. As seen from the Figure, the RP series is quite persistent and is generally below zero. Figure 4 suggests that the jumps and the level of IV are important factors determining the variance risk premium. To investigate formally this conjecture I compute the covariance between RP and past values of TV and JV. In Table 2 I report the results from Wald tests for zero covariance, which confirm the conjectured strong dependence between RP and past values of TV and JV. As seen from equations (11) and (12), this means that the variance risk premium has time-variation which depends on the level of the variance of the continuous price component $\sigma^2(t)$ but importantly also on the past price jumps.

3 Modeling and Inference for Time-Varying Variance Risk Premium

The main question I try to answer in this section is whether we can "rationalize" the empirical evidence of Section 2 for the time-variation in the variance risk premium.

Can we find prices for the different risks in the SV model, which are consistent with no arbitrage and support the empirical findings in Section 2? Are standard prices of risk used in the literature sufficient for this? I start by deriving a pricing kernel that nests the ones commonly used. Following that, I compare the dynamics of the variance risk premium with that implied by the pricing kernel.

3.1 Prices of Risk and Change of Measure

The fundamental theorem of asset pricing implies, under some technical conditions, that no arbitrage is equivalent to the existence of a risk-neutral measure also referred to as Equivalent Martingale Measure (hereafter abbreviated as EMM) under which the discounted gain process associated with an asset is a local martingale. The futures contract involves no initial payment and as a result, assuming that the contract is continuously marked to market, the futures price $F(t)$ is a local martingale under the EMM, see e.g. Duffie (2001).[10] Turning to the specification of the EMM, the presence of jumps in the futures price renders the market essentially incomplete. That is, we cannot complete it by including in the investor's portfolio a finite number of securities[11] which means that in general we have infinitely many EMM-s consistent with no arbitrage. The set of EMM-s is derived in the following theorem.

Theorem 1 *Consider the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with filtration $\mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}$ on which the stochastic volatility model (1)-(4) is defined. Assume that $\mathbf{F}$ is generated by Brownian motions (including B_1 and B_2) and Lévy jumps (including those in the price and the variance). Define a probability measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$ such that $B_1(t) - \int_0^t \psi_1(s)ds$ and $B_2(t) - \int_0^t \psi_2(s)ds$ are two Brownian motions and μ has a compensator $\nu^{\mathbb{Q}}(\omega, dt, dx) := Y(t, \mathbf{x})dtG(d\mathbf{x})$, where $\psi_1 = (\psi_1(t))$, $\psi_2 = (\psi_2(t))$ are predictable processes and $Y = Y(t, \mathbf{x})$ is a strictly positive and predictable function satisfying the following conditions*

$$\int_0^t \psi_i^2(s)ds < \infty \qquad d\mathbb{P} \otimes dt\text{-}a.s \quad and \quad d\mathbb{Q} \otimes dt\text{-}a.s, \quad i = 1, 2, \tag{15}$$

$$\int_0^t \int_{\mathbb{R}_0^n} \left(\sqrt{Y(s, \mathbf{x})} - 1\right)^2 \nu(ds, d\mathbf{x}) < \infty \qquad d\mathbb{P} \otimes dt\text{-}a.s \quad and \quad d\mathbb{Q} \otimes dt\text{-}a.s. \tag{16}$$

$$\int_0^t \int_{\mathbb{R}_0^n} |(Y(s, \mathbf{x}) - 1)\, h(\mathbf{x})|\, dsG(d\mathbf{x}) < \infty, \qquad d\mathbb{P} \otimes dt\text{-}a.s. \tag{17}$$

Then the measure $\mathbb{Q}$ belongs to the set of equivalent martingale measures iff the following condition holds

$$\begin{aligned} b(t) \;\equiv\; & -\frac{1}{2}\sigma^2(t) - \rho\sigma(t)\psi_1(t) - \sqrt{1-\rho^2}\sigma(t)\psi_2(t) \\ & - \int_{\mathbb{R}_0^n} \left(Y(t, \mathbf{x})(e^{h(\mathbf{x})} - 1) - h(\mathbf{x})\right) G(d\mathbf{x}), \qquad d\mathbb{P} \otimes dt\text{-}a.s. \end{aligned} \tag{18}$$

[10] This is subject to a boundedness condition on the interest rate process, but this assumption can be relaxed; see Pozdnyakov and Steele (2004).

[11] Except trivial cases, e.g. when the set of possible jump sizes is finite; see Cont and Tankov (2004).

The stochastic processes $\psi_1(t)$, $\psi_2(t)$ and the (stochastic) function $Y(t, \mathbf{x})$ determine the prices of the different risks in the SV model (1)-(4). $\psi_1(t)$ and $\psi_2(t)$ are the prices for the diffusion type risk in the price and the variance, whereas $Y(t, \mathbf{x})$ determines the compensation for the presence of jumps in the price and the variance. Intuitively, going from physical to risk-neutral distribution, the Brownian/diffusive risks acquire drifts reflecting the compensation for them, while the intensity/compensator for jumps changes by making some jumps more likely to occur than others. $\psi_1(t)$ and $Y(t, \mathbf{x})$ determine together the variance risk premium, which we are after. $\psi_1(t)$ and $\psi_2(t)$ determine the compensation for the diffusive price risk. Since this paper is not interested in the latter, I leave $\psi_2(t)$ unspecified.

The variance risk premium contains compensation for diffusive and jump type risk, since the model has price jumps and in addition $\sigma^2(t)$ contains both a diffusive and a jump factor. The pricing of the diffusive risk is well studied, while pricing of jump risk is less so. Before analyzing different specifications for $\psi_1(t)$ and $Y(t, \mathbf{x})$ I briefly discuss the pricing of jump risk and how it differs from pricing diffusive risk.

The stochastic function $Y(t, \mathbf{x})$ specifies compensation for each possible jump size $\mathbf{x}$ at each point of time t. This is fundamentally different from the pricing of diffusive risk, where at each point of time we have a single price, e.g. $\psi_1(t)$ for the Brownian motion $B_1(t)$. This explains why the market is in general incomplete in the presence of jumps. When there are only diffusive risks we need to include in the portfolio a finite number of instruments (i.e. assets and different derivatives written on them) which have sensitivity towards those diffusive risks and this completes the market. Intuitively, the diffusive risks have a local Gaussian behavior and appropriately weighted set of instruments could completely eliminate (hedge) them. The situation is very different in the presence of jumps where we need a hedging instrument for each possible jump size at each point in time. Thus, provided the jumps have an infinite number of possible jump sizes, the market cannot be completed by a finite number of instruments.

I turn now to modeling the variance risk premium, i.e. modeling $\psi_1(t)$ and $Y(t, \mathbf{x})$ in the SV model (1)-(4). The typical way of specifying the prices of risk, mainly for reasons of tractability, is such that the model is of the same class under both measures (physical and risk-neutral). For the model here this means that under the risk-neutral measure the jumps in the price and the variance are again Lévy processes, and the stochastic variance is a sum of square-root process and jump-driven OU process, possibly with different parameters. This, however, is too restrictive particularly for the jumps. Therefore, I consider also measure changes for jumps which go beyond the "structure-preserving" ones. It is convenient for the subsequent analysis to split the variance risk premium as follows

$$VR_a(t) = VR_a^c(t) + VR_a^j(t),$$

where

$$VR_a^c(t) = \frac{1}{a}\mathbb{E}^{\mathbb{P}}\left(\int_t^{t+a} V^c(s)ds \middle| \mathscr{F}_t\right) - \frac{1}{a}\mathbb{E}^{\mathbb{Q}}\left(\int_t^{t+a} V^c(s)ds \middle| \mathscr{F}_t\right), \tag{19}$$

$$
\begin{aligned}
VR_a^j(t) &= \frac{1}{a}\mathbb{E}^{\mathbb{P}}\left(\int_t^{t+a}\int_{\mathbb{R}_0^n} h^2(\mathbf{x})\mu(ds,d\mathbf{x}) + \int_t^{t+a} V^j(s)ds\middle|\mathscr{F}_t\right) \\
&\quad -\frac{1}{a}\mathbb{E}^{\mathbb{Q}}\left(\int_t^{t+a}\int_{\mathbb{R}_0^n} h^2(\mathbf{x})\mu(ds,d\mathbf{x}) + \int_t^{t+a} V^j(s)ds\middle|\mathscr{F}_t\right).
\end{aligned}
$$

$VR_a^c(t)$ is the part of the variance risk premium which is due to the compensation for the time-variation in the continuous variance factor $V^c(t)$. It is determined by $\psi_1(t)$. The other component of the variance risk premium, $VR_a^j(t)$, is determined only by the price of jump risk $Y(t,\mathbf{x})$. It consists of compensation for the time-variation in $V^j(t)$ as well as a compensation for the presence of jumps in the price. I refer to $VR_a^c(t)$ as diffusive variance risk premium and to $VR_a^j(t)$ as jump variance risk premium.

3.1.1 Specification of $\psi_1(t)$

The price of diffusive risk I consider is specified as follows.
D.$\psi_1(t) = \frac{\lambda_0+\lambda_1 V^c(t)}{\sigma_{cv}\sqrt{V^c(t)}}$, *where* λ_0 *and* λ_1 *are constants such that* $\lambda_0 \geq \frac{\sigma_{cv}^2}{2} - \kappa_c\overline{V}^c \geq 0$.

In the appendix I show that this is a valid price of risk, i.e. Theorem 1 holds. This specification for $\psi_1(t)$ is called extended affine price of risk in Cheridito et al. (2007). It is the most general specification of $\psi_1(t)$ under which $V^c(t)$ is a square-root process under both measures. For this specification of $\psi_1(t)$ we have $VR_a^c(t) = const_0 + const_1 \times V^c(t)$, i.e. the diffusive variance risk premium is an affine function of the diffusive variance factor. This specification implies that the only relevant information at a given time for the diffusive variance risk premium is the level of the diffusive variance factor itself. This change of measure, with the restriction $\lambda_0 \equiv 0$ imposed, is one of the most frequently used for empirical finance applications.

3.1.2 Specification of $Y(t,\mathbf{x})$

The price of jump risk analyzed in the paper is given in the following.
J.$Y(t,\mathbf{x}) = \vartheta_0(\mathbf{x}) + \vartheta_1(\mathbf{x})\tau(t-)$, *where* $\vartheta_0(\mathbf{x}) \geq 0$ *and* $\vartheta_1(\mathbf{x}) \geq 0$ *and in addition*

$$
d\tau(t) = \rho_\tau\tau(t)dt + \int_{\mathbb{R}_0^n}\zeta(\mathbf{x})\mu(dt,d\mathbf{x}), \quad \zeta(\mathbf{x}) \geq 0, \quad \rho_\tau < 0,
$$

$$
\int_{\mathbb{R}_0^n}\zeta(\mathbf{x})G(d\mathbf{x}) < \infty \quad \text{and} \quad \int_{\mathbb{R}_0^n}\zeta^4(\mathbf{x})G(d\mathbf{x}) < \infty,
$$

$$
\int_{\mathbb{R}_0^n}(\vartheta_0(\mathbf{x})-1)^2 G(d\mathbf{x}) < \infty, \quad \int_{\mathbb{R}_0^n}\vartheta_1^2(\mathbf{x})G(d\mathbf{x}) < \infty \quad \text{and} \quad \int_{\mathbb{R}_0^n}\zeta(\mathbf{x})\vartheta_1(\mathbf{x})G(d\mathbf{x}) < -\rho_\tau.
$$

The proof that **J** is a valid price of risk can be found in the appendix. The specification for the price of jump risk implies $VR_a^j(t) = const_0 + const_1 \times \tau(t)$, i.e. $VR_a^j(t)$ is linear in $\tau(t)$. The state variable $\tau(t)$ is modelled as a jump-driven OU process (like $V^j(t)$ but with possibly different parameters (and driving jumps)). There are two special cases of interest. First, if $\vartheta_1 \equiv 0$, the price of risk depends only on the jump size and therefore there is no time variation in the jump risk premium in this case. Given the empirical evidence in the previous section such specification seems implausible. A second special case of the price of risk $\mathbf{J}$ is when $\rho_\tau \equiv \rho_j$ and $\zeta(\mathbf{x}) \equiv k(\mathbf{x})$. In this case we have $\tau(t) \equiv V^j(t)$. The completely general case generalizes the previous two special cases in two directions. First, by allowing $\zeta(\cdot)$ to differ from $k(\cdot)$ I allow for different information, besides the one contained in the jump variance factor $V^j(t)$, to enter the state variable $\tau(t)$. This can be done by setting $\zeta(\mathbf{x})$ to depend on elements in the vector $\mathbf{x}$ on which the function $k(\mathbf{x})$, determining the jumps in the variance, does not depend. The second generalization is to allow the persistence in the state variable $\tau(t)$ to differ from that of the jump variance component $V^j(t)$. That is, $\tau(t)$ and $V^j(t)$ might have the same information (i.e. the jumps in the variance) but this information could be synthesized in a different way. This difference in persistence can be achieved by letting $\rho_\tau \neq \rho_j$.

3.2 Inference for Time-Varying Variance Risk Premium

The prices of risk from the previous subsection imply $VR_a(t) = const_0 + const_1 \times V^c(t) + const_2 \times \tau(t)$, i.e. they determine the time variation in the risk premium. On the other hand, using the VIX index in Section 2 we estimated the variance risk premium without assuming anything about the pricing kernel. Here I match the moments of the estimated variance risk premium with those implied by the pricing kernel and this links the observed variance risk premium with the pricing kernel that generated it. Following the analysis in Section 2, the moments of the variance risk premium that are matched are its mean and its covariance with the past continuous and discontinuous QV. Given the prices of risk it is easy to compute

$$\mathbb{E}\left(VR_a(t)\right) = K_0, \tag{20}$$

$$\mathrm{Cov}\left(VR_a(t), \int_{t-i-a}^{t-i} \sigma^2(s)ds\right) = K_\psi^c e^{-\kappa_c i} + K_\tau^c e^{\rho_\tau i}, \tag{21}$$

$$\mathrm{Cov}\left(VR_a(t), \int_{t-i-a}^{t-i} \int_{\mathbb{R}_0^n} h^2(\mathbf{x})\mu(ds, d\mathbf{x})\right) = K_\tau^j e^{\rho_\tau i}, \tag{22}$$

where K_0, K_ψ^c, K_τ^c and K_τ^j are some constants depending on the magnitude of the compensation for diffusive risk and jumps of different size and sign; K_τ^c is proportional to $\int_{\mathbb{R}_0^n} k(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x})$ and K_τ^j is proportional to $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x})$. Given our interest in the dynamics of the variance risk premium and the semiparametric estimation approach adopted here (i.e. I do not specify a parametric distribution for the jumps) these parameters are all we need. As mentioned in Section 1.2, the main advantage of being semiparametric is the robustness and generality of our estimation results.

Our estimate for the variance risk premium in Section 2 contains estimation error, since it depends on the parameter estimates controlling the SV model, and to avoid errors-in-variables problem I conduct the estimation in one step. That is, all moments including the ones used to estimate the SV model are matched together in a GMM.

Thus, for each value of the parameters I first calculate the linear projection of $IV_a(t)$ on past TV and JV and then construct an estimate for the variance risk premium for this parameter. Using this estimate of the variance risk premium series I evaluate the GMM objective function.

The moments used in this joint estimation are: the ratio of the mean of QV to the mean of the daily squared returns; mean, variance and autocovariance of IV; mean and variance of QV; mean of FV; mean of VR; covariance between VR and past IV; covariance between VR and past QV-IV. The autocovariances of IV used in the estimation are for lags 1, 3 and 6 as well as the average autocovariance for lags $11-20$ and $21-30$.[12] For the covariance between VR and past IV I use the average one for lags of TV $1-10$, $11-20$, $21-30$ and $41-50$. The same is done for the covariance between VR and past values of QV-IV. Thus, overall, I match 20 moments.

The parameters that are estimated are the ones controlling the SV model (given in Section 1.2) as well as K_0, K_ψ^c, K_τ^c, K_τ^j and ρ_τ defined in (20)-(22), which altogether makes 14 parameters to be estimated. To help the interpretation of the estimation results below I briefly elaborate on the implications which different special cases for the prices of risk have on the value of the coefficients in (21)-(22).

- Constant variance risk premium. In this case $K_\tau^c \equiv K_\tau^j \equiv K_\psi^c \equiv 0$. Note that this scenario is observationally equivalent (when the covariances in (21) and (22) are used for identification of the time-variation in the variance risk premium) to the case when $\psi(t)\sqrt{V^c(t)}$ and $\tau(t)$ are time-varying, but are both (at least linearly) independent from $\sigma^2(t)$ and the squared price jumps.[13]

- Affine-in-variance factors variance risk premium. This case was already discussed in Section 2. In this case $\rho_\tau \equiv \rho_j$. Therefore, if the variance jump factor $V^j(t)$ has short memory, this will imply that the impact of the jumps on the variance risk premium dies out quickly as well. Also, since $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x}) > 0$, such specification for the jump risk implies that we can either have $K_\tau^c \equiv K_\tau^j \equiv 0$ or $K_\tau^c \neq 0$ and $K_\tau^j \neq 0$.

- Jump variance risk premium depends only on component of price jumps orthogonal to the variance jumps. In this case $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x}) > 0$ and $\int_{\mathbb{R}_0^n} k(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x}) = 0$. Therefore this scenario implies $K_\tau^c \equiv 0$ and $K_\tau^j \neq 0$.

- Jump variance risk premium depends only on component of variance jumps that is orthogonal to price jumps. In this case $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x}) = 0$ and $\int_{\mathbb{R}_0^n} k(\mathbf{x})\zeta(\mathbf{x})G(d\mathbf{x}) > 0$. The implication of this is $K_\tau^j \equiv 0$ and $K_\tau^c \neq 0$.

[12] In comparison with the estimation in Section 1.2 I dropped one of the moment conditions identifying the persistence in $\sigma^2(t)$. This is done to keep the number of overidentifying restrictions relatively low, following Monte Carlo results in Andersen and Sørensen (1996) on the finite sample performance of GMM.

[13] For the diffusive price of risk **D** this is not possible. For the price of jump risk **J** this is possible. Example is the case where $\mathbf{x} = (x_1, x_2, x_3) > 0$ with $\int_{\mathbb{R}_0^n} 1_{(x_1 x_2 x_3 \neq 0)} G(d\mathbf{x}) = 0$ and $h(\mathbf{x}) = x_1$, $k(\mathbf{x}) = x_2$ and $\zeta(\mathbf{x}) = x_3$.

The estimation results are reported in Table 3. The test for overidentifying restrictions shows that the specification of the prices of risk, together with the stochastic volatility model, provides relatively good fit to the moments used in the estimation. Comparing the estimation results in Table 1 and Table 3, we can see that the parameters determining the evolution of the S&P 500 index under the physical measure do not change substantially when the additional moments identifying the time-variation in the variance risk premium are included in the estimation. Also, note that K_0 is very high in magnitude and statistically significant, confirming the observation made already in Section 2 of a non-trivial variance risk premium. I continue with analysis of the parameters controlling the temporal dependence in the variance risk premium. I summarize the findings in the following points.

(1) The parameters determining the covariance between RP and past values of TV (K^c_ψ and K^c_τ) are not very accurately estimated, as indicated by their relatively big standard errors, and are individually statistically insignificant.

(2) I test the hypothesis that the jumps in the price and/or variance do not determine the time-variation in the variance risk premium. This is equivalent to testing $K^c_\tau \equiv K^j_\tau \equiv 0$. Note that ρ_τ is present only under the alternative hypothesis and is not identified under the null hypothesis. I use a criterion difference test, which has χ^2_2 limiting distribution. The value of the test is 15.55 with corresponding p-value of 0.0004. This indicates that the realized jumps are an important state variable determining the time-variation in the variance risk premium. In addition the test provides also very strong evidence against constant compensation for jump risk, i.e. specification of the form $Y(t, \mathbf{x}) = \vartheta_0(\mathbf{x})$, which has been used in previous work.

(3) Given the strong evidence for the importance of the jumps in determining the time-variation in the variance risk premium, it is interesting to investigate what component of the jumps determines this time-variation. If this is a component in the jumps common for both price and variance we must have $K^c_\tau \neq 0$ and $K^j_\tau \neq 0$. If this is a component contained only in the price jumps (i.e. orthogonal to the variance jumps) we should have $K^c_\tau \equiv 0$ and $K^j_\tau \neq 0$. Similarly, if it is a component present only in the variance jumps we should have $K^c_\tau \neq 0$ and $K^j_\tau \equiv 0$. The coefficient K^j_τ is statistically different from zero, while K^c_τ is not. Therefore, this could be interpreted as evidence that the time-variation in the variance risk premium is determined by component in the price jumps, which is orthogonal to (i.e. independent from) the variance jumps. How can we explain this intuitively? We can link the price jumps with a release of information and associate an increased volatility with disagreement among investors about the effect of this information on the market. Then a price a jump that does not change volatility can be associated with news whose impact on the market is uniformly agreed upon among investors. Thus the results suggest that it is exactly those events that lead to an increase in the risk premia. However, this hypothesis can be true only if the diffusive variance risk, i.e a risk from a persistent shift in the volatility, is also priced.[14] Thus, overall, I conclude that there are two possible scenarios. The first is that the diffusive variance factor does not determine

[14] A Wald test for the hypothesis $K^c_\psi = K^c_\tau = 0$ has a value of 10.99 with a corresponding p-value of 0.0041. Thus, this hypothesis is strongly rejected.

the time-variation in the variance risk premium. In this case the time-variation in the variance risk premium is determined by components of both price and variance jumps. The second possible scenario is that the diffusive variance factor determines the variation in the variance risk premium and in addition a jump component only present in the price jumps determines the variation in the variance risk premium.
(4) The autoregressive coefficients ρ_j and ρ_τ differ significantly. I calculate a Wald test for the hypothesis $\rho_j \equiv \rho_\tau$. The value of the test is 33.77 with a corresponding p-value of 0.0000. In other words the difference between ρ_j and ρ_τ is statistically big. Note that affine-in-variance factors variance risk premium implies $\rho_j \equiv \rho_\tau$. Therefore the estimation results reject strongly such specification.

Overall, the empirical results uncover a non-trivial variance risk premium, whose time-variation depends strongly on the price jumps and the stochastic variance. However, the results here are not conclusive which component of $\sigma^2(t)$, the diffusive (i.e. persistent) or the jump (i.e transient) one, drives the variation in $VR_a(t)$. Further, we can reconcile the variance risk premium variation, but only with the use of general (and quite flexible) specification for the price of jump risk.

The empirical finding of the dependence of the risk premium on past market jumps indicate that investors fear jumps more straight after big stock market shocks and are willing to pay higher premium to protect from such events. There are at least two possible economic explanations for this effect. The first is that straight after big market changes investors view the occurrence of jumps more likely. Such reassessment of the jump probability can naturally occur in the context of a Bayesian investor. The reason is that the parameters controlling jumps are hard to estimate precisely and a single realization of a big jump can tilt the posterior distribution of the parameters towards values implying more frequent jump activity.

A second explanation for the persistent effect of the jumps on the risk premium is habit persistence in investor's fear of jumps. This can be generated in a habit persistence type equilibrium model (as in Campbell and Cochrane (1999) for example) in which habits are affected by jumps. In such a model the representative investor will treat differently the diffusive and jump risk, as in Liu, Pan, and Wang (2005) and Bates (2008). When a jump occurs investor's willingness to protect her portfolio against future big (negative) jumps increases.

On the practical side, the increased fear of future jumps straight after a market crash means that the protection for them will be more expansive. This means higher prices for close-to-maturity deep-out-of-the-money options as they form a natural hedge for jumps. I now explicitly use this observation to provide further support for the pricing of jump risk.

4 Nonparametric Robustness Checks

In our analysis of the variance risk premium, and in particular its dynamics, we found that jumps play a crucial role. On one hand, jumps are present in the returns and the stochastic volatility, where there have short term effect, while on the other hand jumps have persistent effect on the variance risk premium through increased fear of future jumps. In this section I provide nonparametric evidence for both these claims.

4.1 Evidence for Jumps in Returns and Volatility

It is easy to detect the presence of jumps in the returns using the 5-minute high-frequency S&P 500 index returns data. In Figure 5, I plot the histogram of the jump test statistic of Ait-Sahalia and Jacod (2009) over all of the days in the sample.[15] A value of 1 of the test statistics indicates the presence of jumps and a value of 2 indicates no jumps. As seen from the figure there are a lot of days in the sample with jumps. Although the results in this paper are derived for an arbitrary jump measure μ (see equation (1)), the nonparametric evidence in Figure 5 suggests that μ is more likely of infinite activity, or at least of a very high-intensity finite activity jump process.[16] This is in line with the results in Bakshi et al. (2008), who report evidence in favor of infinite activity jumps based on a very flexible parametric jump specification.

Although stochastic volatility is unobserved, I provide nonparametric evidence for presence of jumps in volatility using the VIX index. Jumps in the stochastic volatility will imply jumps in the VIX index when the jump volatility factor has some persistence (like in our semiparametric specification). I apply the procedure of Lee and Mykland (2008) to test for jumps in the VIX index on each day in the sample.[17] Following recommendation in Lee and Mykland (2008), I use window size of 16 days in computing the local volatility of the differenced VIX index series and select a time horizon of 22 days in setting up the rejection region (the test applies for a fixed time horizon). For significance level of 5%, I find a total of 65 days in which the VIX index jumps. The median absolute value of the jump size is 2.71 (in annualized volatility units).[18] Further, I applied the same test for the daily returns on the S&P 500 index and found that on 20 of these 65 days there was an associated jump in the S&P 500 index.

To provide further evidence for cojumping in the returns and the stochastic volatility, following Jacod and Todorov (2009), I calculate "realized" correlations between squared jumps in the S&P 500 and VIX series over periods of 22 days. The estimated median correlation is 0.7, while if there was no common jumping this number should be close to 0. Further, I compute the share in the cojump variation measure of Jacod and Todorov (2009) due to negative jumps in the returns and positive jumps in the volatility (which arrive at the same time). I find that 63% of the cojump variation in the sample is due to this combination of jumps. These findings are in line with

[15] The test of Ait-Sahalia and Jacod (2009) is derived for arbitrary, finite or infinite activity, jumps.

[16] There are also a lot of days in the sample where the test statistic takes values somewhere in the middle of the interval $(1, 2)$ (or even outside this interval). This reflects the statistical uncertainty associated with the test - for sharper results clearly much higher frequencies are needed, e.g. as the ones used in Ait-Sahalia and Jacod (2009).

[17] I would like to thank a referee for suggesting this.

[18] These results however should be interpreted with some caution for two reasons. First, the frequency of the VIX data is relatively low and therefore a lot of jumps are probably missed by the test. Second, the test is designed to estimate Poisson-type jumps, while the model here (1)-(4) allows for far more active jumps.

the strong significance of the estimated parameter $\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x})$ reported in Tables 1 and 3. We can further link the nonparametric evidence here with the results in Bakshi et al. (2008), who find that only the negative jumps in the returns carry a risk premium. The nonparametric evidence here suggests that in most cases those jumps are accompanied by positive jumps in volatility and therefore the latter are also being priced.

In sum, the nonparametric evidence presented in this section strongly supports the jump structure of the SV model.

4.2 Realized Jumps and Risk Premia

The key implication of the price of jump risk $\mathbf{J}$ is that past "realized" jumps should have an impact on the intensity of the future ones *under the risk-neutral measure*. To check this in a completely non-parametric setting, I use a result that follows easily from Carr and Wu (2003) (see the proof of their Proposition 1). It is based on the intuitive fact that short maturity deep out-of-the-money put option prices are mainly determined by the price jumps. More formally, the result is the following. Let $P_t(K,T)$ denote the price at time t of a put with strike K expiring at time T and $k = \ln(K/F_t)$ is the moneyness (where F_t is the futures price with expiration at T). Then we have

$$IJ_t(k) := \lim_{T\downarrow t} \frac{P_t(T,K)}{F_t(T-t)} = \int_{-\infty}^{k} \int_{\mathbb{R}_0^n} e^l 1_{[h(\mathbf{x})<l]} \nu_t^{\mathbb{Q}}(d\mathbf{x})dl, \quad k<0, \tag{23}$$

where $\nu_t^{\mathbb{Q}}(d\mathbf{x})$ is the "local" intensity of the jumps under the risk-neutral measure.[19] For the price of jump risk $\mathbf{J}$ we have

$$IJ_t(k) = \int_{-\infty}^{k} \int_{\mathbb{R}_0^n} e^l 1_{[h(\mathbf{x})<l]} (\nu_0(\mathbf{x}) + \nu_1(\mathbf{x})\tau(t))G(d\mathbf{x})dl, \quad k<0, \tag{24}$$

which suggests *that under* $\mathbf{J}$, $IJ_t(k)$ should be an affine function of the state variable $\tau(t)$ and thus depend on the past "realized" jumps. To investigate this implication of the price of jump risk, I collected from Option Metrics short maturity put option prices for the period January 4, 1996 till November 29, 2002 (Option Metrics provides data for the period after 1996 only). For each day I use the closest to maturity put option closing mid-quote to proxy $IJ_t(k)$[20] (options with maturities at least 9 days were considered only). Following Carr and Wu (2003), I set $k = -9.2\%$[21] and linearly interpolate the put option price for that level of moneyness on each of the days. On Figure 6 I plot the correlation between $IJ_t(k)$ and the past squared realized jumps. As seen from the figure, and consistent with the modeling of jump risk in this paper, the intensity of the price jumps under the risk-neutral measure depends persistently

[19]That is, the jump compensator of the jumps under the risk-neutral measure is $\nu_t^{\mathbb{Q}}(d\mathbf{x})dt$.

[20]The simulation results in Carr and Wu (2003) suggest that for maturities less than a month (which are used here) this is a reasonable approximation.

[21]Experiments with other levels of moneyness produced similar results and are available upon request.

on the past "realized" jumps. Note that a model where jump intensity is linear in the stochastic volatility as in Bates (2000) and Pan (2002) can generate empirically realistic dependence between the risk-neutral jump intensity and the past realized jumps only if the squared jumps were a persistent process *under the measure* $\mathbb{P}$, which is not supported by the data (as seen from Figure 2). Thus, the nonparametric results in this subsection provide strong support for risk-neutral jump specification in which past jumps affect their future intensity.

5 Conclusion

This paper provides an arbitrage-free explanation of the market variance risk premium dynamics in the framework of a semiparametric stochastic volatility model. Jumps play key role in explaining the observed risk premium. On one hand their occurrence in the market level is documented and it is typically linked with a spike in the volatility which dies out quickly. On the other hand the occurrence of jumps leads to a persistent increase in the variance risk premium. These two empirical findings are rationalized with a price of jump risk that increases after big market moves which is suggestive of changing attitude of investors towards jumps.

There are several important directions in which the current work can be extended. First, a natural next step of the analysis is to find an equilibrium explanation for the estimated dynamics of the pricing kernel. Second, the pricing of jump risk can be further analyzed using separately deep out-of-the money put and call prices. This can show whether the effect of extreme events on the market is different for positive and negative jumps. Third, the analysis here can be performed on individual stock level and the linkages with the market variance risk premium analyzed.

Appendix: Equivalent Martingale Measures

A Proof of Theorem 1

We start with a Lemma, which is of independent interest.

Lemma 1 *Consider the probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *with filtration* $\mathbf{F} = (\mathscr{F}_t)_{t\in\mathbb{R}_+}$. *Suppose that the filtration* $\mathbf{F}$ *is generated by* d*-dimensional standard Brownian motion* W *and* n*-dimensional homogenous Poisson measure* μ *with compensator* $\nu^{\mathbb{P}}$ *(under the measure* $\mathbb{P}$*). Let* ψ *be a* $d \times 1$ *predictable process and* $Y(\omega, t, \mathbf{x})$ *a strictly positive and predictable function. Denote with* $\mathbb{Q}$ *a probability measure under which* $W(t) - \int_0^t \psi(s)ds$ *is a standard Brownian motion and the random measure* μ *has compensator* $\nu^{\mathbb{Q}}(\omega, dt, d\mathbf{x}) = Y(\omega, t, \mathbf{x})\nu^{\mathbb{P}}(\omega, dt, d\mathbf{x})$ *(assuming that such a measure exists!). Assume that* $\mathbb{P}_0 \sim \mathbb{Q}_0$ *and in addition the following conditions are satisfied*

$$\int_0^t \psi(s)\psi^{'}(s)ds < \infty \qquad d\mathbb{P} \otimes dt\text{-}a.s \quad and \quad d\mathbb{Q} \otimes dt\text{-}a.s, \tag{A.1}$$

$$\int_0^t \int_{\mathbb{R}_0^n} \left(\sqrt{Y}-1\right)^2 \nu(ds, d\mathbf{x}) < \infty \qquad d\mathbb{P}\otimes dt\text{-}a.s \quad and \quad d\mathbb{Q}\otimes dt\text{-}a.s. \tag{A.2}$$

Then we have $\mathbb{P} \overset{loc}{\sim} \mathbb{Q}$*, that is for every finite* $t > 0$ *we have* $\mathbb{P}_t \sim \mathbb{Q}_t$*.*

Proof. Under the measure $\mathbb{P}$ a representation theorem holds and therefore this measure is unique (note we are in the canonical setting because of the assumption for the filtration). Furthermore since the characteristics of (W, μ) are constant under $\mathbb{P}$ we have even *local uniqueness* (definition III.2.37 in Jacod and Shiryaev (2003)) for this probability measure (this follows from Theorem III.2.40 in Jacod and Shiryaev (2003)). Define $T := \inf(t : H(t) = \infty)$, where

$$H(t) = \int_0^t \psi(s)\psi^{'}(s)ds + \int_0^t \int_{\mathbb{R}_0^n} \frac{\left(\sqrt{Y}-1\right)^2}{Y} \nu^{\mathbb{Q}}(ds, d\mathbf{x}).$$

Consider the process

$$Z^{'}(t) = \begin{cases} \mathcal{E}\left(-\int_0^t \psi(s)dW^{\mathbb{Q}}(s) + \int_0^t \int_{\mathbb{R}_0^n} \left(\frac{1}{Y}-1\right)\tilde{\mu}^{\mathbb{Q}}(ds, d\mathbf{x})\right) & \text{for } t < T \\ 0 \qquad \text{for } t \geq T, & \end{cases}$$

where $\tilde{\mu}^{\mathbb{Q}}(\omega, dt, dx) := \mu(\omega, ds, d\mathbf{x}) - \nu^{\mathbb{Q}}(\omega, dt, dx)$ and $\mathcal{E}$ stands for a stochastic exponential (see Jacod and Shiryaev (2003) for a definition). Taking into account the relationship $\nu^{\mathbb{Q}}(dt, d\mathbf{x}) = Y(\omega, t, \mathbf{x})\nu(dt, d\mathbf{x})$ and using the conditions (A.1) and (A.2) we have $H(t) < \infty \qquad d\mathbb{P}\otimes dt$-a.s.

Combining everything we can apply Theorem III.5.34 in Jacod and Shiryaev (2003) (adapted to the case when the filtration is generated by a d-dimensional Brownian motion and n-dimensional homogenous Poisson measure) to conclude $\mathbb{P} \overset{loc}{\ll} \mathbb{Q}$ with density process $Z^{'}$. Therefore, to prove the (local) equivalence of the two measures we need only to show that $\mathbb{Q}(Z^{'}(t) = 0) = 0$. This is an easy consequence of the conditions (A.1) and (A.2). As a result we have $\mathbb{Q} \overset{loc}{\ll} \mathbb{P}$. Using Theorem III.5.19 (adapted to the case when the filtration is generated by a d-dimensional Brownian motion and n-dimensional homogenous Poisson measure) in Jacod and Shiryaev (2003) the density process of $\mathbb{Q}$ with respect to $\mathbb{P}$ is

$$Z(t) = \begin{cases} \mathcal{E}\left(\int_0^t \psi(s)dW^{\mathbb{Q}}(s) + \int_0^t \int_{\mathbb{R}_0^n} (Y-1)\,\tilde{\mu}(ds, d\mathbf{x})\right) & \text{for } t < T \\ 0 \qquad \text{for } t \geq T. & \end{cases}$$

□

From this Lemma and the conditions in Theorem 1 it follows that $\mathbb{P}$ and $\mathbb{Q}$ are locally equivalent. Therefore we are left with showing that $F(t)$ is a local martingale under $\mathbb{Q}$ iff condition (18) holds. The condition in (17) guarantees that the quadratic covariation process $\left[\int_0^t \int_{\mathbb{R}_0^n} h(\mathbf{x})\tilde{\mu}(ds, d\mathbf{x}), Z(t)\right]$ has locally integrable variation under $\mathbb{P}$ and therefore the following is a well defined local martingale under $\mathbb{Q}$

$$\int_0^t \int_{\mathbb{R}_0^n} h(\mathbf{x})\tilde{\mu}^{\mathbb{Q}}(ds, d\mathbf{x}) = \int_0^t \int_{\mathbb{R}_0^n} h(\mathbf{x})\tilde{\mu}(ds, d\mathbf{x}) - \int_0^t \int_{\mathbb{R}_0^n} h(\mathbf{x})(Y-1)\nu(ds, d\mathbf{x}).$$

Based on that, $f(t)$ satisfies the following SDE under the measure $\mathbb{Q}$

$$\begin{aligned} df(t) &= \left(b(t) + \rho\psi_1(t)\sigma(t) + \sqrt{1-\rho^2}\psi_2(t)\sigma(t)\right) dt + \int_{\mathbb{R}_0^n} \left(Y(\omega, t, \mathbf{x}) - 1\right) h(\mathbf{x})\nu(dt, d\mathbf{x}) \\ &\quad +\sigma(t)\left(\rho dB_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2} dB_2^{\mathbb{Q}}(t)\right) + \int_{\mathbb{R}_0^n} h(\mathbf{x})\tilde{\mu}^{\mathbb{Q}}(dt, d\mathbf{x}). \end{aligned}$$

Using Ito's lemma we have for $F(t)$ under the measure $\mathbb{Q}$

$$\begin{aligned} \frac{dF(t)}{F(t-)} &= \left(b(t) + \frac{1}{2}\rho\psi_1(t)\sigma(t) + \sqrt{1-\rho^2}\psi_2(t)\sigma(t)\right) dt \\ &\quad + \int_{\mathbb{R}_0^n} \left(Y(\omega, t, \mathbf{x})(e^{h(\mathbf{x})} - 1) - h(\mathbf{x})\right)\nu(ds, d\mathbf{x}) \\ &\quad +\sigma(t)\left(\rho dB_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2} dB_2^{\mathbb{Q}}(t)\right) + \int_{\mathbb{R}_0^n} \left(e^{h(\mathbf{x})} - h(\mathbf{x}) - 1\right)\tilde{\mu}^{\mathbb{Q}}(dt, d\mathbf{x}). \end{aligned}$$

Since $F(t)$ must be a local martingale under the measure $\mathbb{Q}$ we need to set the drift term in the SDE above to zero. From here we get the result in (18). $\square$

B Proof for the Diffusive Risk Price $\psi_1(t)$

We need to show that specification **D** for $\psi_1(t)$ generates an equivalent change of measure (i.e. that condition (A.1) is satisfied). The process V^c is a square-root process under both measures. The dynamics of $V^c(t)$ under the measure $\mathbb{Q}$ is given by

$$dV^c(t) = \left(\lambda_0 + \kappa_c\overline{V}^c + (\lambda_1 - \kappa_c)V^c(t)\right) dt + \sigma_{cv}\sqrt{V^c(t)} dB_1^{\mathbb{Q}}(t).$$

If $\kappa_c\overline{V}^c \geq 0$ and $\lambda_0 \geq 0$ the square-root process (under both measures) satisfies the Yamada-Watanabe condition and therefore has a unique non-explosive solution under both measures. This implies that for the equivalence of the measures $\mathbb{P}$ and $\mathbb{Q}$ we need only verify that $V^c(t) > 0 \quad d\mathbb{P} \otimes dt - a.s.$ and $\mathbb{Q} \otimes dt - a.s.$ To check this condition we need to analyze the behavior of V^c at the boundary 0. The necessary and sufficient conditions for non-attainment of the boundary under the measure $\mathbb{P}$ and $\mathbb{Q}$ respectively, starting from a strictly positive value, are: $\sigma_{cv}^2 \leq 2\kappa_c\overline{V}^c$ and $\sigma_{cv}^2 \leq 2\kappa_c\overline{V}^c + 2\lambda_0$ (see Ikeda and Watanabe (1981) for example, these conditions guarantee that the boundary is entrance under both measures, i.e. starting from a positive value it is never reached in finite time and if the process starts from zero it always goes out). Therefore, for those values of the parameters condition (A.1) is satisfied and hence we have equivalence of the two measures.

C Proof for the Jump Price of Risk $Y(t,\mathbf{x})$

First, it is easy to derive that for the specification of $Y(\omega,t,\mathbf{x})$ we have

$$\int_0^t \int_{\mathbb{R}_0^n} \left(\sqrt{Y}-1\right)^2 \nu(ds,d\mathbf{x}) \le 2t \int_{\mathbb{R}_0^n} \left(\vartheta_0(\mathbf{x})-1\right)^2 G(d\mathbf{x}) + 2\int_0^t \tau(s)ds \int_{\mathbb{R}_0^n} \vartheta_1^2(\mathbf{x})G(d\mathbf{x}),$$

therefore it sufficient for condition (A.2) to hold that the following is true

$$\mathbb{E}^{\mathbb{P}}\left(\int_0^t \tau(s)ds\right) < \infty \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}\left(\int_0^t \tau(s)ds\right) < \infty \quad \text{for } \forall t > 0.$$

The condition is trivially satisfied under the measure $\mathbb{P}$. I will show that it holds under the measure $\mathbb{Q}$ as well. Note that

$$\int_{\mathbb{R}_0^n} \zeta(\mathbf{x})\vartheta_0(\mathbf{x})G(d\mathbf{x}) \le \sqrt{\int_{\mathbb{R}_0^n} \zeta^2(\mathbf{x})G(d\mathbf{x}) \int_{\mathbb{R}_0^n} \left(\vartheta_0(\mathbf{x})-1\right)^2 G(d\mathbf{x})} + \int_{\mathbb{R}_0^n} \zeta(\mathbf{x})G(d\mathbf{x}) < \infty,$$

$$\int_{\mathbb{R}_0^n} \zeta(\mathbf{x})\vartheta_1(\mathbf{x})G(d\mathbf{x}) \le \sqrt{\int_{\mathbb{R}_0^n} \zeta^2(\mathbf{x})G(d\mathbf{x}) \int_{\mathbb{R}_0^n} \vartheta_1^2(\mathbf{x})G(d\mathbf{x})} < \infty.$$

Similar inequalities hold true when $\zeta(\mathbf{x})$ is replaced with $\zeta^2(\mathbf{x})$. Therefore, the claim follows from the following general result.

Lemma 2 (a) *There exists probability measure on the canonical probability space such that the canonical process V is a semimartingale with initial distribution $\mathcal{L}(V(0)) = \eta$ (with positive support) and satisfies the following equation*

$$dV(t) = \rho V(t)dt + \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(dt,d\mathbf{x}), \tag{C.1}$$

where $k : \mathbb{R}_0^n \to \mathbb{R}_+$, μ is integer-valued measure on $\mathbb{R}_+ \times \mathbb{R}_0^n$ with compensator $\nu(ds,d\mathbf{x}) = ds\left(m(V(s-))G_1(d\mathbf{x}) + G_2(d\mathbf{x})\right)$ where $m : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function, $G_1 : \mathbb{R}_0^n \to \mathbb{R}_+$, $G_2 : \mathbb{R}_0^n \to \mathbb{R}_+$ and

$$m(x) \le C \vee x, \quad \textit{where } C \textit{ is some constant},$$

$$K_1 := \int_{\mathbb{R}_0^n} k(\mathbf{x})G_1(d\mathbf{x}) < \infty, \quad K_2 := \int_{\mathbb{R}_0^n} k(\mathbf{x})G_2(d\mathbf{x}) < \infty,$$

$$K_1' := \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G_1(d\mathbf{x}) < \infty, \quad K_2' := \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G_2(d\mathbf{x}) < \infty.$$

(b) *In addition to the conditions in part (a) assume that $-\rho > K_1 \ge 0$ and $m(x) = x$. Then $V(t)$ is asymptotically covariance stationary.*

Proof.

Part (a) The characteristics of the process V with truncation function $h(x) = x$ (that is without truncation) are given by

$$B(t) = \int_0^t (K_1 m(V(s-)) + \rho V(s-) + K_2)\, ds,$$

$$\widetilde{C}(t) = \int_0^t (K_1' m(V(s-)) + K_2')\, ds,$$

$$\nu(ds, A) = ds \int_{\mathbb{R}_0^n} 1_{[k(\mathbf{x})\in A)]} \left(m(V(s-))G_1(d\mathbf{x}) + G_2(d\mathbf{x})\right), \quad A \in \mathbb{R}_0.$$

See Jacod and Shiryaev (2003), chapter II.2, for a definition of the characteristics of a general semimartingale. $\widetilde{C}(t)$ stands for the second modified characteristic. I define a sequence of semimartingales $(\widetilde{V}_K)$ with initial distribution η and the following characteristic triplet (the truncation function is again $h(x) = x$)

$$B_K(t) = \int_0^t \left(K_1 \left(m(\widetilde{V}_K(s-)) \wedge K\right) + \rho \left(\widetilde{V}_K(s-) \wedge K\right) + K_2\right) ds,$$

$$\widetilde{C}_K(t) = \int_0^t \left(K_1' \left(m(\widetilde{V}_K(s-)) \wedge K\right) + K_2'\right) ds,$$

$$\nu_K(ds, A) = ds \int_{\mathbb{R}_0^n} 1_{[k(\mathbf{x})\in A)]} \left(\left(m(\widetilde{V}_K(s-)) \wedge K\right) G_1(d\mathbf{x}) + G_2(d\mathbf{x})\right), \quad A \in \mathbb{R}_0.$$

I will show that such processes exist. First, for each $K > 0$ the characteristics of the semimartingale $\widetilde{V}_K$ are majorized, i.e.

$$\sup_{\omega,t} |K_1 \left(m(\widetilde{V}_K(s-)) \wedge K\right) + \rho \left(\widetilde{V}_K(t-) \wedge K\right) + K_2| < \infty,$$

$$\sup_{\omega,t} |K_1' \left(m(\widetilde{V}_K(t-)) \wedge K\right) + K_2'| < \infty.$$

Also, $\left(K_1' \left(m(\widetilde{V}_K(s)) \wedge K\right) + K_2'\right)$ and $\left(K_1 \left(m(\widetilde{V}_K(s-)) \wedge K\right) + \rho \left(\widetilde{V}_K(s) \wedge K\right) + K_2\right)$ are continuous in $\widetilde{V}_K(s)$. This holds true also for

$$\left(\left(m(\widetilde{V}_K(s)) \wedge K\right) \int_{\mathbb{R}_0^n} g(\mathbf{x})G_1(d\mathbf{x}) + \int_{\mathbb{R}_0^n} g(\mathbf{x})G_2(d\mathbf{x})\right),$$

for all continuous and bounded functions $g(\mathbf{x})$ vanishing around zero. Finally, since $K_1 < \infty$ and $K_2 < \infty$, we trivially have

$$\lim_{a\uparrow\infty} \sup_{\omega} \left(\left(m(\widetilde{V}_K(t-)) \wedge K\right) \int_{\mathbb{R}_0^n} 1_{|k(\mathbf{x})|>a} G_1(d\mathbf{x}) + \int_{\mathbb{R}_0^n} 1_{|k(\mathbf{x})|>a} G_2(d\mathbf{x})\right) = 0, \quad \text{for } \forall t \geq 0.$$

Therefore, the conditions of Theorem IX.2.31 in Jacod and Shiryaev (2003) are satisfied. This implies that there exists probability measure (denoted hereafter with $\mathbb{P}^K$) supporting $\widetilde{V}_K$ (the canonical process) as a semimartingale with characteristics $(B_K, \widetilde{C}_K, \nu_K)$ and initial distribution η. I will now show that the sequence of processes $(\widetilde{V}_K)$ converges weakly (upon taking a subsequence if necessary) to a limiting process and will identify the limit with the process V. To establish weak convergence I prove that the sequence $(\widetilde{V}_K)$ is tight. For this I use Theorem VI.4.18 in Jacod and Shiryaev (2003). It is sufficient to show that

1. For all $T > 0$ and $\epsilon > 0$

$$\lim_{a\uparrow\infty} \limsup_{K} \mathbb{P}^K[\nu_K([0,T] \times \{\mathbf{x} : |k(\mathbf{x})| > a\}) > \epsilon] = 0. \tag{C.2}$$

2. The following sequence of processes is C-tight (i.e. the sequence of processes is tight and all its limit points are continuous processes)

$$\begin{aligned} F_K(t) &= \int_0^t \left(K_1\left(m(\widetilde{V}_K(s-)) \wedge K\right) + |\rho|\left(\widetilde{V}_K(s-) \wedge K\right) + K_2\right) ds \\ &\quad + \int_0^t \left(K_1'\left(m(\widetilde{V}_K(s-)) \wedge K\right) + K_2'\right) ds. \end{aligned}$$

First I establish that (F_K) is C-tight. For this I make use of Theorem VI.3.26 in Jacod and Shiryaev (2003). Note that $F_K(t)$ is absolutely continuous. Therefore for the C-tightness of $F_K(t)$ it suffices to show that the process $\widetilde{V}_K(t)$ satisfies the following boundedness in probability condition

$$\lim_{a\uparrow\infty} \sup_{K} \mathbb{P}^K\left[\sup_{0\le s\le t} \widetilde{V}_K(s) > a\right] = 0, \quad \text{for } \forall t \geq 0.$$

We have $\widetilde{V}_K(t) = \widetilde{V}_K(0) + \rho\int_0^t \left(\widetilde{V}_K(s-) \wedge K\right) ds + \int_0^t \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(ds, d\mathbf{x})$, therefore $\widetilde{V}_K(t) \le \widetilde{V}_K(0) + |\rho|\int_0^t \left(\widetilde{V}_K(s-) \wedge K\right) ds + \int_0^t \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(ds, d\mathbf{x})$, and since $\int_0^t \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(ds, d\mathbf{x}) \geq 0$ as $k(\mathbf{x}) > 0$, and $\widetilde{V}_K(0) > 0$ as η has a positive support, using Gronwall's inequality (see Revuz and Yor (1994) for example) we have

$$\widetilde{V}_K(s) \le \left(\widetilde{V}_K(0) + \int_0^t \int_{\mathbb{R}_0^n} k(\mathbf{x})\mu(ds, d\mathbf{x})\right) \exp(|\rho|t), \quad \text{for } 0 \le s \le t.$$

Therefore the C-tightness of (F_K) will be established if we can show that $\mathbb{E}^K\left(\widetilde{V}_K(t)\right) < C$, where C is a constant that does not depend on K (the compensator for the jumps is given by ν_K). We have

$$\mathbb{E}^K\left(\widetilde{V}_K(t)\right) = \mathbb{E}\left(\widetilde{V}_K(0)\right) + \rho\int_0^t \mathbb{E}^K\left(\widetilde{V}_K(s-) \wedge K\right) ds + K_1\int_0^t \mathbb{E}^K\left(m(\widetilde{V}_K(s-)) \wedge K\right) ds + tK_2,$$

and since $m(x) \le x \vee C$ for some constant C, for $K > C$ we have

$$\mathbb{E}^K\left(\widetilde{V}_K(t)\right) \le \mathbb{E}\left(\widetilde{V}_K(0)\right) + \rho \int_0^t \mathbb{E}^K\left(\widetilde{V}_K(s-) \wedge K\right) ds + K_1 \int_0^t \mathbb{E}^K\left(\widetilde{V}_K(s-) \wedge K\right) ds + t(K_2 + C).$$

If $K_1 + \rho \le 0$ we have $\mathbb{E}^K\left(\widetilde{V}_K(t)\right) \le \mathbb{E}\left(\widetilde{V}_K(0)\right) + t(K_2 + C)$, and note that the right hand side of the above inequality does not depend on K. If $K_1 + \rho > 0$ we have

$$\mathbb{E}^K\left(\widetilde{V}_K(t)\right) \le \mathbb{E}\left(\widetilde{V}_K(0)\right) + (K_1 + \rho) \int_0^t \mathbb{E}^K\left(\widetilde{V}_K(s)\right) ds + t(K_2 + C),$$

therefore using Gronwall's inequality we have

$$\mathbb{E}^K\left(\widetilde{V}_K(t)\right) \le \left(\mathbb{E}\left(\widetilde{V}_K(0)\right) + T(K_2 + C)\right) \exp\left((K_1 + \rho)\, t\right), \quad 0 \le t \le T,$$

and note again that the right hand side of the above inequality does not depend on K.

This proves C-tightness of the sequence of processes (F_K). To establish tightness of the sequence $(\widetilde{V}_K)$ we need only verify that condition (C.2) holds. We have

$$\begin{aligned}
\mathbb{P}^K[\nu_K([0,T] \times \{\mathbf{x} : |k(\mathbf{x})| > a\}) > \epsilon] &\le \frac{\mathbb{E}^K\left(\nu_K([0,T] \times \{\mathbf{x} : |k(\mathbf{x})| > a\})\right)}{\epsilon} \\
&\le \frac{\int_0^T \mathbb{E}^K\left(m(\widetilde{V}_K(s)) \wedge K\right) ds \int_{\mathbb{R}_0^n} 1_{|k(\mathbf{x})|>a} G_1(d\mathbf{x}) + \int_{\mathbb{R}_0^n} 1_{|k(\mathbf{x})|>a} G_2(d\mathbf{x})}{\epsilon} \\
&\le \frac{C \int_{\mathbb{R}_0^n} k(\mathbf{x}) G_1(d\mathbf{x}) + \int_{\mathbb{R}_0^n} k(\mathbf{x}) G_2(d\mathbf{x})}{a\epsilon},
\end{aligned}$$

where C is a constant that does not depend on K. For the last inequality I made use of the result derived above that $\mathbb{E}^K(\widetilde{V}_K(t))$ is bounded by a constant, which does not depend on K. This proves that the sequence $(\widetilde{V}_K)$ is tight.

Now we are left with identifying the limiting process with the process V. For this I use Theorem IX.2.22 in Jacod and Shiryaev (2003). It suffices to establish the following

$$\int_0^t |\widetilde{V}_K(s) \wedge K - \widetilde{V}_K(s)| ds \xrightarrow{p} 0, \quad \text{for every } \forall t > 0, \text{ as } K \uparrow \infty,$$

$$\int_0^t |m(\widetilde{V}_K(s)) \wedge K - m(\widetilde{V}_K(s))| ds \xrightarrow{p} 0, \quad \text{for every } \forall t > 0, \text{ as } K \uparrow \infty.$$

The first result follows since for arbitrary $s > 0$ and $\epsilon > 0$ we have

$$\mathbb{P}^K\left(|\widetilde{V}_K(s) \wedge K - \widetilde{V}_K(s)| > \epsilon\right) \le \mathbb{P}^K\left(\widetilde{V}_K(s) > K\right) \le \frac{\mathbb{E}^K\left(\widetilde{V}_K(s)\right)}{K},$$

and as shown above $\mathbb{E}^K(\widetilde{V}_K(s))$ can be bounded by a constant, which does not depend on K. The second result follows analogously since

$$\begin{aligned}
&\mathbb{P}^K\left(|m(\widetilde{V}_K(s)) \wedge K - m(\widetilde{V}_K(s))| > \epsilon\right) \leq \mathbb{P}^K\left(m(\widetilde{V}_K(s)) > K\right) \\
&\leq \frac{\mathbb{E}^K\left(m(\widetilde{V}_K(s))\right)}{K} \leq \frac{\mathbb{E}^K\left(\widetilde{V}_K(s)\right) + C}{K} \qquad .
\end{aligned}$$

Part (b) We can write

$$\mathbb{E}\left(V(t)|\mathscr{F}_s\right) = e^{\rho(t-s)}V(s) + K_1 \int_s^t e^{\rho(t-u)}\mathbb{E}\left(V(u)|\mathscr{F}_s\right))du + K_2 \int_u^t e^{\rho(t-u)}du, \quad t \geq s.$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}\left(V(t)|\mathscr{F}_0\right) &= e^{(K_1+\rho)t}V(0) + K_2 \int_0^t e^{(K_1+\rho)(t-u)}du \\
&= e^{(K_1+\rho)t}V(0) - \frac{K_2}{\rho + K_1}\left(1 - e^{(K_1+\rho)t}\right),
\end{aligned}$$

and since $-\rho - K_1 > 0$ we have

$$\lim_{t\to\infty} \mathbb{E}\left(V(t)|\mathscr{F}_0\right) = -\frac{K_2}{\rho + K_1},$$

that is we have asymptotic stationarity in the mean. For the second moment we have

$$\mathbb{E}\left(V^2(t)|\mathscr{F}_0\right) = \left(\mathbb{E}\left(V(t)|\mathscr{F}_0\right)\right)^2 + K_1' \int_0^t e^{2\rho(t-u)}\mathbb{E}\left(V(u)|\mathscr{F}_0\right)du + K_2' \int_0^t e^{2\rho(t-u)}du.$$

Using the expression for $\mathbb{E}\left(V(t)|\mathscr{F}_s\right)$ and after simplifying and taking the limit as $t \to \infty$ we have

$$\lim_{t\to\infty} \mathbb{E}\left(V^2(t)|\mathscr{F}_0\right) = \left(\frac{K_2}{\rho + K_1}\right)^2 + \frac{K_1'}{2\rho}\frac{K_2}{\rho + K_1} - \frac{K_2'}{2\rho},$$

and therefore we have asymptotic stationarity in the second moment. □

References

Ait-Sahalia, Y. and J. Jacod (2009). Testing for Jumps in a Discretely Observed Process. *Annals of Statistics 37*, 184–222.

Alizadeh, S., M. W. Brandt, and F. Diebold (2002). Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance 57*, 1047–1091.

Andersen, T., L. Benzoni, and J. Lund (2002). An Empirical Investigation of Continuous-Time Equity Return Models. *Journal of Finance 57*, 1239–1284.

Andersen, T., T. Bollerslev, and F. Diebold (2005). Parametric and Nonparametric Measurement of Volatility. In Y. Ait-Sahalia and L. Hansen (Eds.), *Handbook of Financial Econometrics*. North-Holland.

Andersen, T. and B. Sørensen (1996). GMM Estimation of a Stochastic Volatility Model: A Monte Casrlo Study. *Journal of Business and Economic Statistics 14*, 328–352.

Andrews, D. (2001). Testing when a Parameter is on the Boundary of the Maintained Hypotheis. *Econometrica 69*, 683–734.

Bakshi, G., P. Carr, and L. Wu (2008). Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economies. *Journal of Financial Economics 87*, 132–156.

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

Bakshi, G., N. Kapadia, and D. Madan (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies 16*, 101–143.

Bakshi, G. and D. Madan (2000). Spanning and Derivative-Security Valuation. *Journal of Financial Economics 55*, 205–238.

Bakshi, G. and D. Madan (2006). A Theory of Volatility Spreads. *Management Science 52*, 1945–1956.

Barndorff-Nielsen, O. E., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales. In Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Springer.

Barndorff-Nielsen, O. E. and N. Shephard (2001). Non- Gaussian Ornstein-Uhlenbeck-based Models and Some of Their Applicaions in Financial Economics. *Journal of the Royal Statistical Society: Series B 63*, 167–241.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics 4*, 1–30.

Bates, D. (2000). Post-'87 Crash Fears in S&P 500 Future Options. *Journal of Econometrics 94*, 181–238.

Bates, D. (2008). The Market for Crash Risk. *Journal of Economic Dynamics and Control 32*, 2291–2321.

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

Bollerslev, T., R. Engle, and D. Nelson (1994). ARCH Models. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4. Amsterdam: North-Holland.

Bollerslev, T., M. Gibson, and H. Zhou (2005). Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. Working paper, Duke University.

Brandt, M. and K. Wang (2003). Time-Varying Risk Aversion and Expacted Inflation. *Journal of Monetary Economics 58*, 1457–1498.

Britten-Jones, M. and A. Neuberger (2000). Option Prices, Implied Price Processes, and Stochastic Volatility. *Journal of Finance 55*, 839–866.

Broadie, M., M. Chernov, and M. Johannes (2007). Model Specification and Risk Premia: Evidence from Futures Options. *Journal of Finance 62*, 1453–1490.

Campbell, J. and J. Cochrane (1999). By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy 107*, 205–251.

Carr, P. and L. Wu (2003). What Type of Process Underlies Options? A Simple Robust Test. *Journal of Finance 58*, 2581–2610.

Carr, P. and L. Wu (2009). Variance Risk Premia. *Review of Financial Studies*, forthcoming.

Cheridito, P., D. Filipović, and R. Kimmel (2007). Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics 83*, 123–170.

Chernov, M., R. Gallant, E. Ghysels, and G. Tauchen (2003). Alternative Models for Stock Price Dynamics. *Journal of Econometrics 116*, 225–257.

Chernov, M. and E. Ghysels (2000). A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk-Neutral Measures for the Purpose of Options Valuation. *Journal of Financial Economics 56*, 407–458.

Cont, R. and P. Tankov (2004). *Financial Modelling With Jump Processes*. London: Chapman & Hall.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton: Princeton University Press.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*, 1343–1376.

Eraker, B. (2004). Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *Journal of Finance 59*, 1367–1403.

Eraker, B., M. Johannes, and N. Polson (2003). The Impact of Jumps in Volatility and Returns. *Journal of Finance 58*, 1269–1300.

Ikeda, N. and S. Watanabe (1981). *Stochastic Differential Equations and Diffusion Processes*. Tokyo: North-Holland.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems For Stochastic Processes* (2nd ed.). Berlin: Springer-Verlag.

Jacod, J. and V. Todorov (2009). Testing for Common Arrivals of Jumps for Discretely Observed Multidimensional Processes. *Annals of Statistics*, forthcoming.

Jiang, G. and R. Oomen (2008). A New Test for Jumps in Asset Prices. *Journal of Econometrics 114*, 352–370.

Lee, S. and P. Mykland (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *Review of Financial Studies 21*, 2535–2563.

Liu, J., J. Pan, and T. Wang (2005). An Equilibrium Model of Rare-Event Premia and Its Implications for Option Smirks. *Review of Financial Studies 18*, 131–164.

Pan, J. (2002). The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study. *Journal of Financial Economics 63*, 3–50.

Pozdnyakov, V. and J. Steele (2004). On the Martingale Framework for Futures Prices. *Stochastic Processes and their Applications 109*, 69–77.

Revuz, D. and M. Yor (1994). *Continuous Martingales and Brownian Motion* (2nd ed.). New York: Springer-Verlag.

Todorov, V. (2009a). Econometric Analysis of Jump-Driven Stochastic Volatility Models. *Journal of Econometrics*, forthcoming.

Todorov, V. (2009b). Estimation of Coninuous-time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics*, forthcoming.

Table 1: Estimation results for the SV model (1)-(4)

Panel A. Parameter Estimates

Parameter	Estimate	Standard Error
π	0.2368	0.0175
η	1.0717	0.0750
κ_c	0.0381	0.0100
σ_c	1.0549	0.1018
$-\rho_j$	1.1760	0.1761
$\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(d\mathbf{x})$	2.5742	0.4472
$\int_{\mathbb{R}_0^n} h^2(\mathbf{x})G(d\mathbf{x})$	0.1998	0.0209
$\int_{\mathbb{R}_0^n} h^4(\mathbf{x})G(d\mathbf{x})$	0.3431	0.1696
$\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x})$	0.8555	0.2101
GMM test of overidentifying restrictions		
χ^2	1.0498	
d.o.f	(3)	
p-value	0.7892	

Panel B. Moment Condition Tests

Moment Condition	t-statistic
autocorrelation in IV for lag 1	-0.1710
autocorrelation in IV for lag 3	0.2799
autocorrelation in IV for lag 6	0.2754
aver. autocorrelation in IV for lags $10-20$	0.6859
aver. autocorrelation in IV for lags $20-30$	0.6492
aver. autocorrelation in IV for lags $30-40$	0.2316
$\mathbb{E}(IV(t))$	-0.3574
$\mathbb{E}(IV^2(t))$	-0.1222
$\mathbb{E}(QV(t) - IV(t))$	0.1217
$\mathbb{E}(QV^2(t))$	0.2436
$\mathbb{E}(FV_\delta^2(t))$	0.7760
$1 - \pi$	1.1817

Note: The model is estimated using GMM-type estimator with moment conditions specified in Section 1.2. The asymptotic variance-covariance matrix, used for calculating the optimal weighting matrix, is computed using Parzen weights with a lag length of 80. In the estimation the following nonnegativity and stationarity restrictions are imposed: $\sigma_c < \eta$, $\kappa_c > 0$ and $\rho_j < 0$.

Table 2: Wald tests for zero covariances between RP and lags of TV and RP and lags of JV

Covariance between RP and lags of TV			Covariance between RP and lags of JV		
Lags	Wald test	P-value	Lags	Wald test	P-value
1	13.6909	0.0002	1	7.7395	0.0054
5	16.4802	0.0056	5	18.8518	0.0021
10	32.2537	0.0004	10	26.8376	0.0028
15	54.2592	0.0000	15	33.6741	0.0038
20	70.5563	0.0000	20	50.7947	0.0002
25	81.9608	0.0000	25	59.6582	0.0001
30	101.353	0.0000	30	71.4084	0.0000

Note: The Wald statistic tests the null hypothesis of zero covariances between RP and lags of TV (respectively JV) up to the corresponding lag. The RP measure was constructed, using the SV model (1)-(4) with parameter values the estimated ones reported in Table 1. For the calculation of the asymptotic variance of the covariances the error from the estimation of the parameters was taken into account.

Table 3: Parameter Estimates from the Joint Estimation

Parameter	Estimate	Standard Error
π	0.2169	0.0161
η	0.9462	0.1023
κ_c	0.0300	0.0130
σ_c	0.8831	0.1636
$-\rho_j$	1.4933	0.2595
$\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(d\mathbf{x})$	2.6087	0.6509
$\int_{\mathbb{R}_0^n} h^2(\mathbf{x})G(d\mathbf{x})$	0.1811	0.0212
$\int_{\mathbb{R}_0^n} h^4(\mathbf{x})G(d\mathbf{x})$	0.3034	0.1128
$\int_{\mathbb{R}_0^n} h^2(\mathbf{x})k(\mathbf{x})G(d\mathbf{x})$	0.7204	0.1976
K_0	-0.4378	0.1435
K_ψ^c	1.1683	2.4335
K_τ^c	-1.4006	2.4246
K_τ^j	-0.0909	0.0349
$-\rho_\tau$	-0.0200	0.0101
GMM test of overidentifying restrictions		
χ^2	6.3136	
d.o.f	(6)	
p-value	0.3890	

Note: The model is estimated using GMM-type estimator with moment conditions specified in Section 3. The asymptotic variance-covariance matrix, used for calculating the optimal weighting matrix, is computed using Parzen weights with a lag length of 80. In the estimation the following nonnegativity and stationarity inequality restrictions are imposed: $\sigma_c < \eta$, $\kappa_c > 0$ and $\rho_j < 0$.

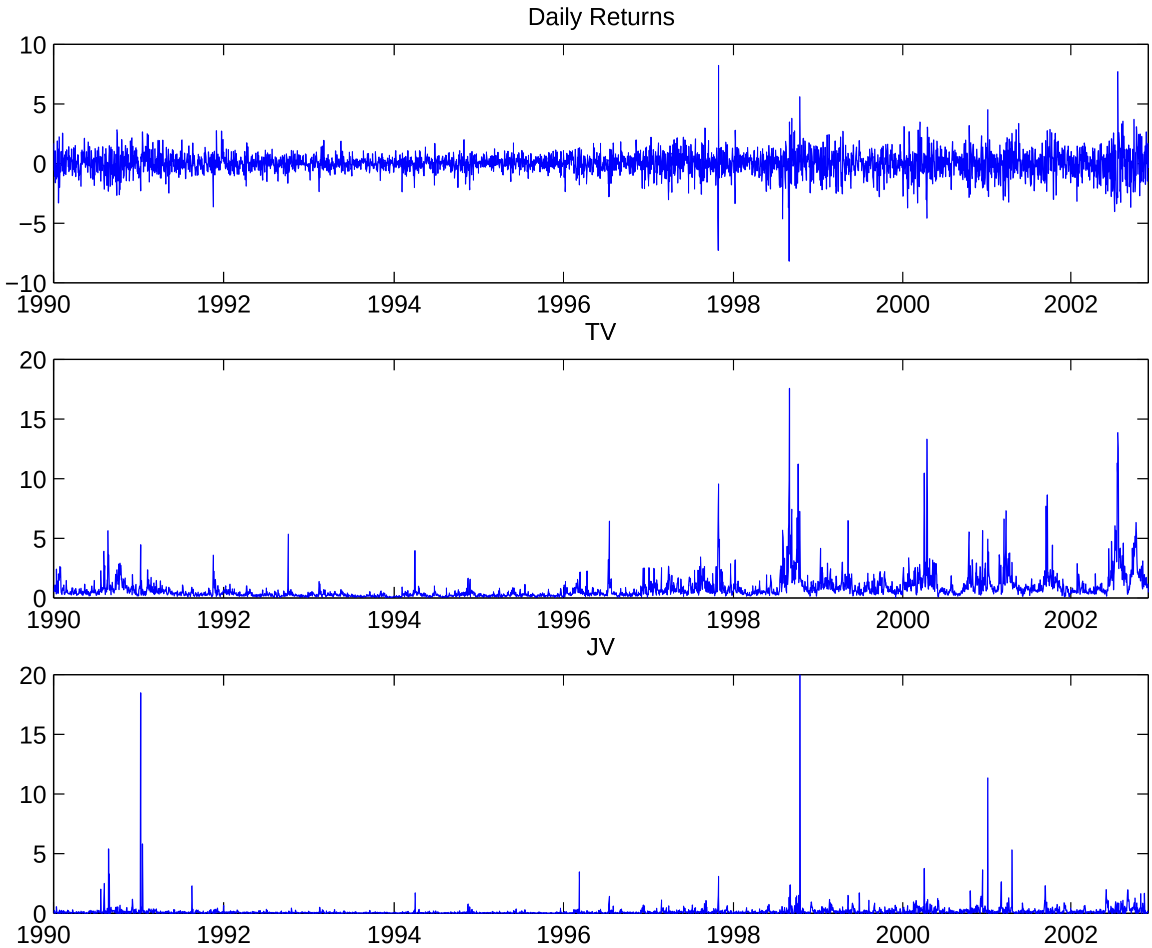

Figure 1: S&P 500 daily measures. The top panel shows daily returns; the middle panel shows the daily TV and the bottom panel shows the daily difference between RV and TV. The sample period is from January 2 1990 till November 29 2002 and includes 3256 daily high-frequency observations on the S&P 500 futures contract.

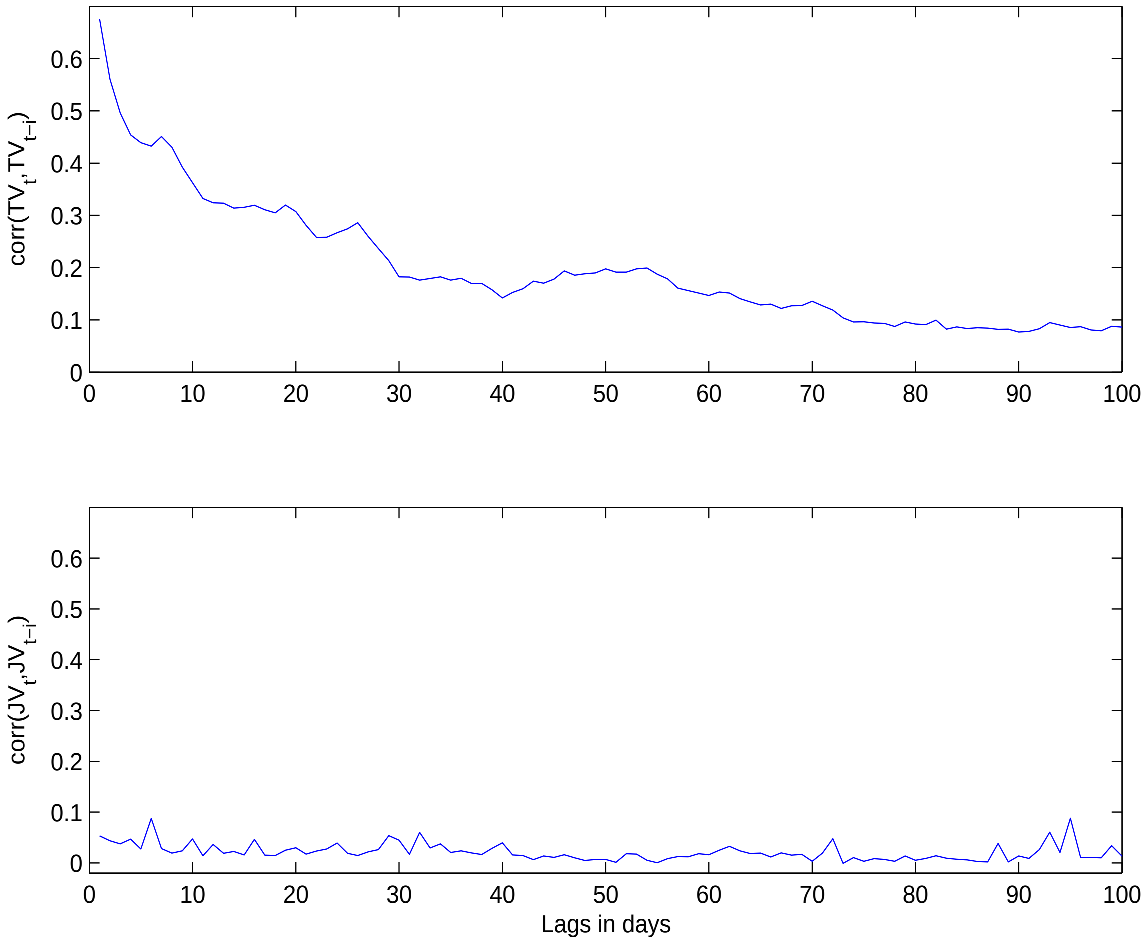

Figure 2: S&P 500 sample correlations. The top panel shows the sample autocorrelation in TV and the second panel shows the autocorrelation in JV=RV-TV.

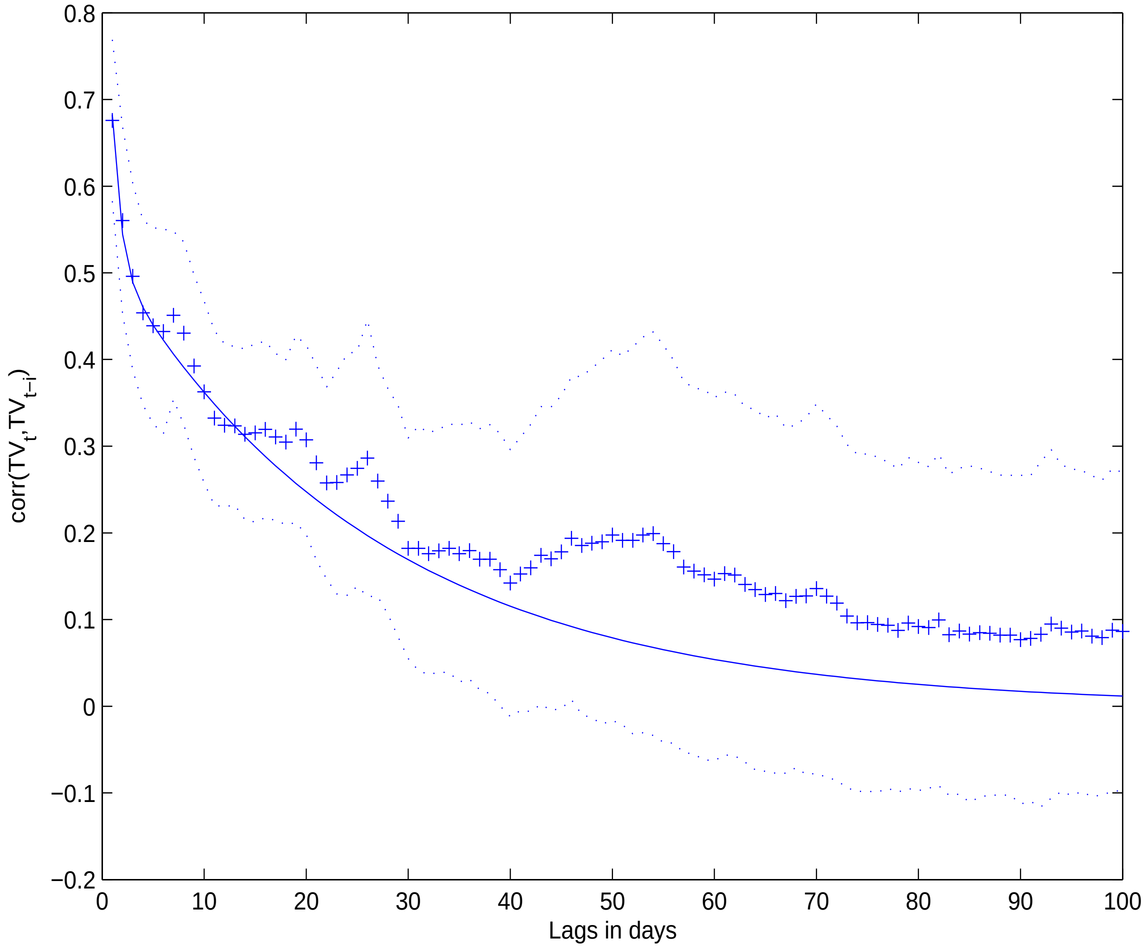

Figure 3: The figure shows the empirical and the fitted autocorrelation for TV. The empirical autocorrelation of the TV is marked with +. The dashed lines are the 95% confidence interval for the autocorrelation with GMM-type standard errors. The solid line is the autocorrelation implied from the SV model given in (1)-(4). The parameters were set at the estimated values reported in Table 1.

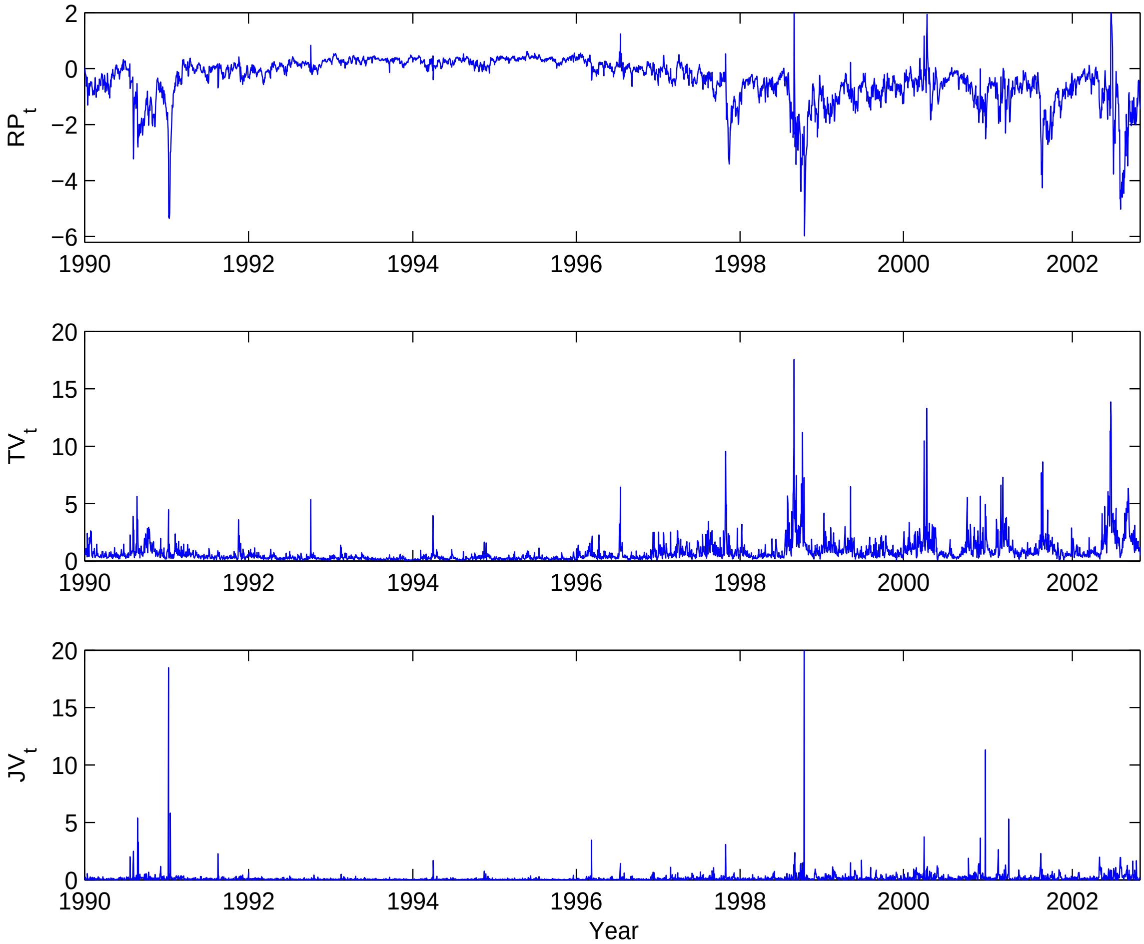

Figure 4: Estimated Variance Risk Premium. The top panel shows the RP measure calculated using equation (10). The middle panel shows the daily TV and the bottom panel shows the daily difference between RV and TV. The sample period is from January 2 1990 till November 29 2002 and includes 3256 daily observations on the VIX index as well as 3256 daily high-frequency observations on the S&P 500 futures contract.

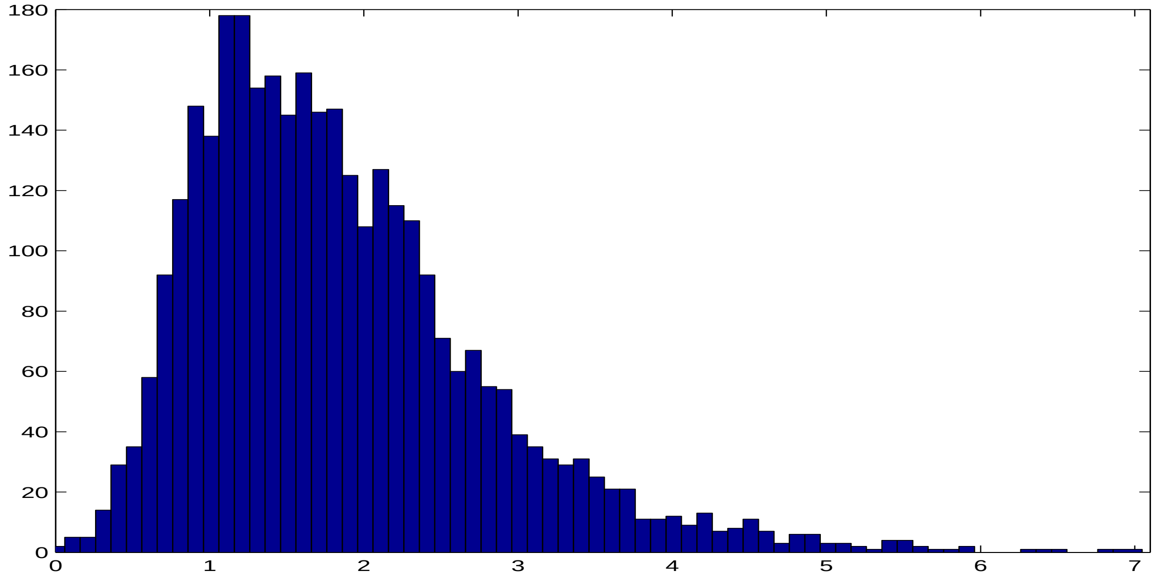

Figure 5: Robustness check for presence of price jumps. The figure shows the histogram of the jump test of Ait-Sahalia and Jacod (2009) computed for each day in the sample using the 5-minute S&P 500 index returns. A value of 1 indicates presence of jumps during the day and a value of 2 indicates no jumps.

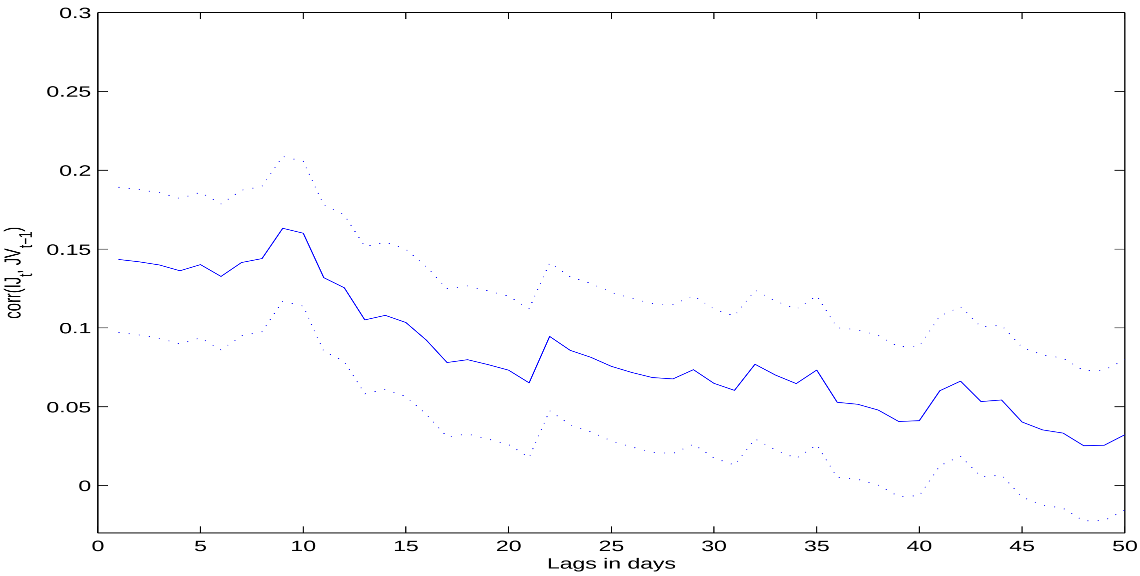

Figure 6: Robustness check for the price of jump risk **J**. The figure shows the estimated correlation between IJ and past JV, where IJ is the measure defined in (24).