Option returns and dynamic risk premia: a direct approach

Piotr Orłowski[*], András Sali[†], Fabio Trojani[‡]

February 14, 2015

Abstract

We propose a computationally tractable estimation approach for a completely specified (under $\mathbb{P}$ and $\mathbb{Q}$) multifactor stochastic volatility models that aims to fit the dynamic properties of returns on option trading strategies. We show in a Monte Carlo experiment that our approach delivers reliable results even under moderate misspecification. We estimate a model using returns on delta-hedged option portfolios as observables. We describe the empirical properties of such returns and recover their model-implied conditional second moment structure. Return-fitted models exhibit lacking pricing properties. Including price information in the estimation significantly worsens the model's ability to plausibly describe delta-hedged option returns.

[*]University of Lugano, Swiss Finance Institute. E-mail: piotr.orlowski@usi.ch
[†]University of Lugano, Swiss Finance Institute. E-mail: andras.sali@usi.ch
[‡]University of Lugano, Swiss Finance Institute

Hunt in every jungle, for there is wisdom and good hunting in all of them. (African proverb)[a]

[a]Martin H. Manser, *The Facts on File Dictionary of Proverbs*

1 Introduction

In a Black-Scholes type world with no priced time-varying risks options would be traded mostly as a means of avoiding the costs associated with the dynamic trading that could perfectly replicate them. In today's equity markets trading the underlying contract (or a futures written upon it) is cheap and becoming cheaper, while equity derivatives markets are burgeoning. Therefore it is justified to believe that options are more than simply a dynamic trading strategy in the underlying asset. In a world of multiple sources of risk (e.g. stochastic volatility and jumps), the replication strategy is not a perfect substitute for the option. The remarkable liquidity of certain option markets hints at the importance of this replication risk, which we shall denote as unspanned risk in this paper[1]. The economically significant premia associated with unspanned risks that have been documented in the literature[2] further strengthen the importance of option markets, and by extension the importance of information embedded in option prices for model estimation.

Derivative prices have been used directly as model inputs in estimation over the course of the last decades. The models were used as *filtering devices* for inference about many types of (unspanned) risk premia. Model outputs were also used for studying the dynamic properties of otherwise unobservable variables which were assumed to drive the risks in the economy. Thus the development of parametric models complemented the development of non-parametric methods over the same period.

[1]These markets are often characterized by very "cheap" trading in the underlying through futures contracts or ETFs

[2]For example, Carr and Wu [2009] demonstrated how to recover the variance risk premium from options and established its basic empirical properties.

In this paper we argue that augmenting the set of model inputs with *returns on traded options* improves the performance and reliability of parametric models as filtering devices for conditional risk premia. A model which fully specifies the statistical ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) dynamics of the asset price and the risk factors imposes tight links on the stochastic discount factor ($d\mathbb{Q}/d\mathbb{P}$). Various types of inputs play different roles in pin-pointing the estimated quantities: derivative prices identify $\mathbb{Q}$ parameters, high-frequency variation measures allow to uncover the $\mathbb{P}$ laws governing the latent processes, but the resulting $d\mathbb{Q}/d\mathbb{P}$ is a pure computational consequence of the mathematical structure of the model. The correctness of the resultant structure for risk premia is a direct consequence of the correctness of the model specification. Our contribution is to introduce *delta-hedged returns on derivatives* as data on the *realizations of premia for unspanned risks* as a direct estimation target for $d\mathbb{Q}/d\mathbb{P}$. Thus we argue that our approach improves the uncovering of the dynamic properties of premia for unspanned risks even in misspecified models.

Concentrating on portoflio as opposed to *individual option* returns brings numerous benefits. Most importantly, the statistical properties of returns on individual out of the money options are ill-suited for standard analysis, as documented by Broadie et al. [2009] (see e.g. figure 1 therein). To the contrary, returns on delta-hedged option portfolios exhibit meaningful variation (see figure 1).

Our main contributions are (i) augmenting the range set of financial data useful for model estimation with delta-hedged derivative portfolio returns (ii) the introduction of a computationally efficient estimation method of multifactor latent affine stochastic volatility models based on delta-hedged option returns; the method identifies all $\mathbb{P}$ and $\mathbb{Q}$ parameters; (iii) introduction of methods for assessing the risk-return profile of option portfolio investments; (iv) assessing the ability of traditional option price estimation to capture conditional risk premia in case of both well and misspecified models (v) evidence on Sharpe ratios available in variance swap trading and contrasting them against other option trades available for investors (such as variance swaps).

Our simulation results suggest that our computational method is accurate for assessing conditional risk premia and risk (Sharpe ratios) on diverse delta-hedged option trading strategies. It is also effective in gen-

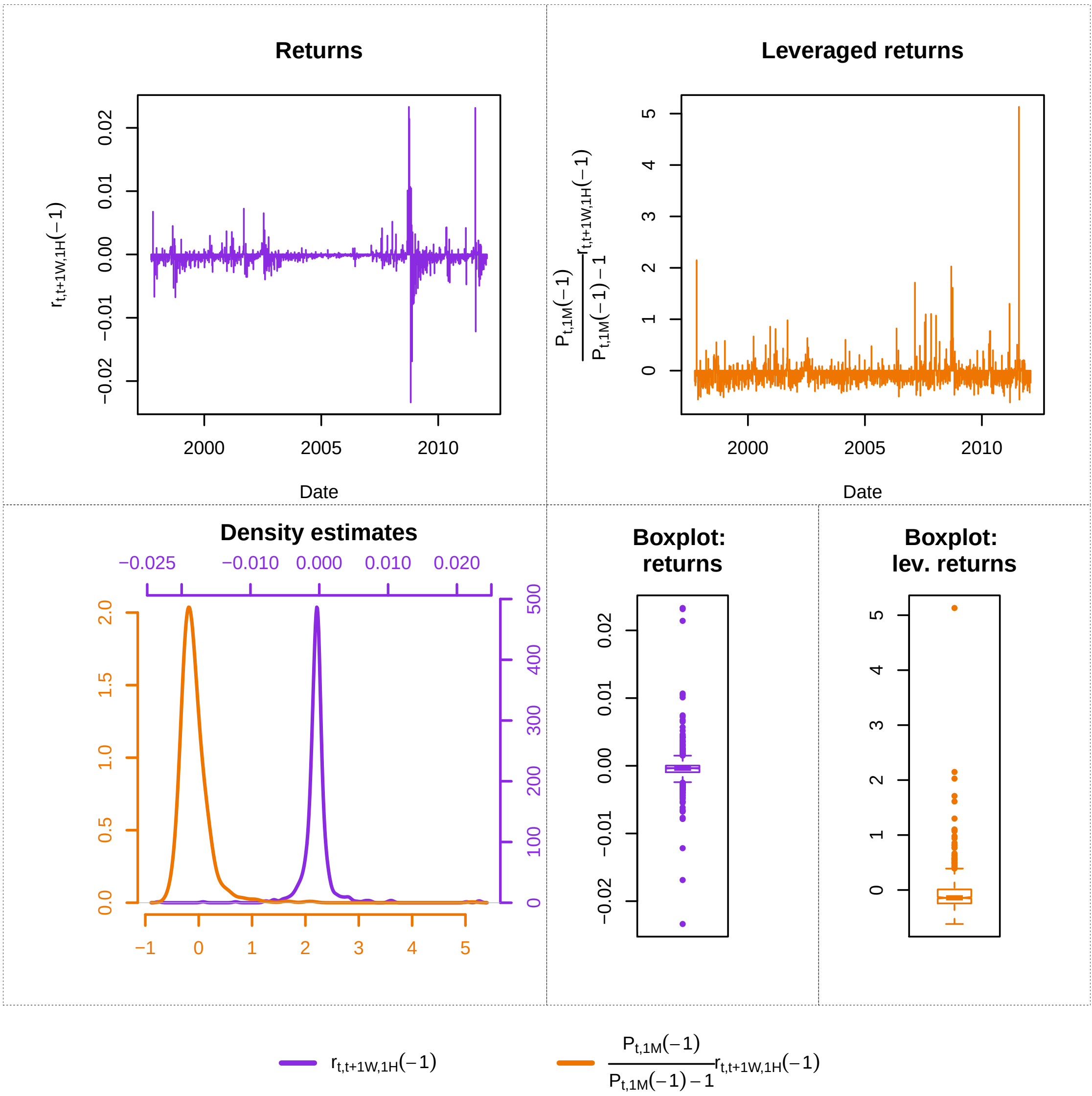


Figure 1: Returns (top left), leveraged returns (top right), kernel density estimates (bottom left) and boxplots (bottom right). The returns are on a portfolio representing the $u = -1$ divergence (defined in section 3.1), delta-hedged every hour. Leveraging is defined as pre-multiplying the return $r_{t,t+h}(u)$ by $\frac{P_{t,t+h}(u)}{P_{t,t+h}(u)-1}$. The sample ranges from 1997-09-24 to 2012-02-01.

erating good trading signals for option portfolios, that outperform statically mean-variance efficient portfolios. We also show through simulations that traditional price based estimation is useful for conditional risk-premium estimation in the well-specified case but underperforms our approach in the misspecified case. Finally, we analyze the conditional unspanned risk premia estimates for the S&P 500 option market.

The paper is organized as follows. In Section 2 we motivate the notion of unspanned risks in option returns and we introduce the problem of estimating conditional risk premia and risk, whilst relating to the existing literature. In Section 3 we introduce the class of power option portfolios, the cornerstone of our analysis. We illustrate their tractability in affine models and analyze their properties. Section 4 introduces a version of the Kálmán filter for the estimation problem. Section 5 studies the performance of our approach compared to traditional price calibration on simulated data, both when the models are well specified and misspecified, section6 presents a preliminary empirical analysis. Section 7 concludes.

2 Compensation for unspanned risks

The Variance swap is the most common example of an unspanned risk contract. It consists of a position in a portfolio of options which are subsequently delta-hedged over the life of the contract. The (forward) price of the flexible leg of such contracts equals the $\mathbb{Q}$-expected quadratic variation of the option underlying over a given period. The variance risk premium (henceforth VRP) is, under the assumptions of a continuous price process, equal to the *expected payoff of a delta-hedged option portfolio.* Indeed if one constructs an option portfolio that pays the logarithm of the terminal stock price and until the termination of the options continuously delta-hedged, then the time t conditional expected annualized payoff is just the time t conditional VRP:

$$VRP_{t,T} = \frac{1}{T-t}\left(\mathbb{E}_t^{\mathbb{P}}(RV_{t,T}) - \mathbb{E}_t^{\mathbb{Q}}(RV_{t,T})\right) . \tag{1}$$

where $RV_{t,T}$ is the realized quadratic variation of the underlying asset. In this sense, the VRP measures the expected value of the tracking error, i.e. the error that cannot be eliminated by hedging infinitely frequently in the underlying market. It is possible to construct many different portfolios whose payoff, after continuous hedging, is dominated by realized quadratic variation. This naturally leads to the characterization of unspanned risks by looking at the tracking error of a class of option portfolios.

The VRP as defined above is the expected *payoff* gained through the acquisition of the replicating option portfolio. When considering the lucrativeness of an investment, it is more natural to consider the *return* on a unit of investment. To this end, we can consider the log risk premia (Carr and Wu [2009]) or similarly the return risk premia:

$$LRP_{t,T} = \log\left(\frac{\mathbb{E}_t^{\mathbb{P}}(RV_{t,T})}{\mathbb{E}_t^{\mathbb{Q}}(RV_{t,T})}\right) \tag{2}$$

$$RRP_{t,T} = \frac{\mathbb{E}_t^{\mathbb{P}}(RV_{t,T})}{\mathbb{E}_t^{\mathbb{Q}}(RV_{t,T})} - 1 \tag{3}$$

While the expected *return* to holding the delta-hedged option portfolio is a reasonable measure of the lucrativeness of "buying" unspanned risk, it is not the best measure of evaluating the compensation for this risk. A broadly used measure in the empirical asset pricing literature is the Sharpe ratio of an investment strategy. In this spirit we can consider the annualized variance Sharpe ratio as:

$$VSR_{t,T} = \sqrt{T-t}\frac{VRP_{t,T}}{\sqrt{\mathbb{V}_t^{\mathbb{P}}(RV_{t,T})}} = \sqrt{T-t}\cdot\frac{RRP_{t,T}}{\sqrt{\mathbb{V}_t^{\mathbb{P}}\left(\frac{RV_{t,T}}{\mathbb{E}_t^{\mathbb{Q}}(RV_{t,T})}\right)}} \tag{4}$$

Whilst time variation in the VSR is arguably useful for assessing a time-varying compensation for unspanned risks, it is generally not observable. Contrary to conditional VRP and RRP which have a conditionally unbiased non-parametric estimator, the conditional VSR estimation requires a model for determining the conditional moments of realized variance.

It is useful to note why price-based estimation is not always appropriate for estimating the dynamics of risk compensation. From Equation (4) it is clearly visible that accurate prices by themselves do not necessarily imply accurate description of the VSR, since the latter depends on the dynamic properties of $RV_{t,T}$. Thus, in order to accurately describe conditional risk compensation, an accurate characterization of the conditional realized variance under the measure $\mathbb{P}$ is needed.

Most estimation approaches do not enforce this constraint as they often abstract from the properties of $RV_{t,T}$. Additionally, in price-based estimation, the implied states and their dynamics are mostly determined by the strict and dominating restrictions that the price panels enforce upon them. In case of misspecification, this can lead to spurious inferred states and RV dynamics, rendering implied risk premia estimates inaccurate.

2.1 An illustration of difficulties

Whilst the VRP, RRP and VSR are all empirically interesting, estimating them accurately is challenging. To illustrate this point, we simulated a 4000-week sample of option price data, using a two-factor model described in detail in section 5. We applied three estimation techniques to the data:

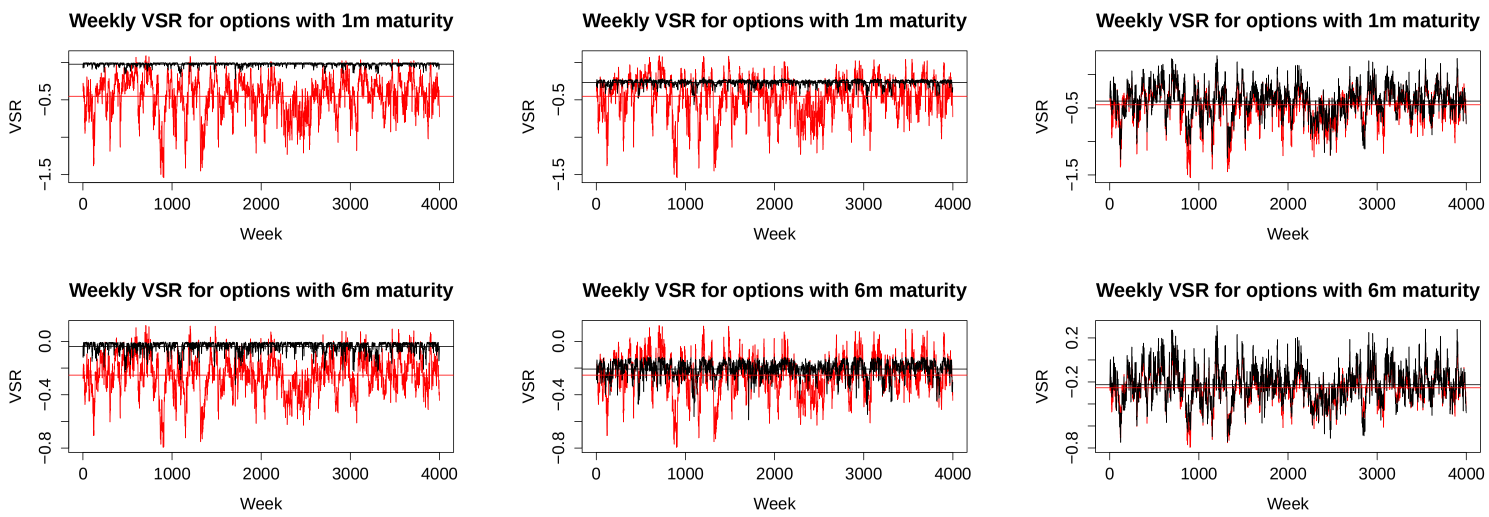


Figure 2: Simulated and estimated VSRs; leftmost panel: price-based estimation of a misspecified model; middle panel: return-based estimation of a misspecified model; right panel: return-based estimation of the true model. Red line represents true conditional weekly Sharpe ratios while black line represents model-based estimates.

1. A Kálmán filter on option prices, as in Gruber et al. [2015]. The model uses no data on the underlying asset, hence $\mathbb{P}$ jump parameters have to be constrained to be equal to $\mathbb{Q}$ jump parameters. The model's pricing kernel is thus misspecified, since the simulated model did exhibit a jump risk premium. Estimated and true Sharpe ratios are plotted in the leftmost panel of figure 2.
2. An estimation based on delta-hedged option portfolio returns and returns on the underlying asset, fully desribed in section 4. The model is left misspecified as above so that its "ability" to fit the data is not changed.
3. The return-based estimation of a correctly specified model, i.e. with different $\mathbb{P}$ and $\mathbb{Q}$ jump distributions.

Figure 2 illustrates the estimated and simulated conditional one-week Sharpe ratios[3], with the panels left to right corresponding to estimation approaches listed above. A number of interesting observations are immediate:

1. Shutting down the jump risk premium channel (leftmost panel) makes it impossible to match the Sharpe ratios on variance swap positions. This is not surprising as a significant part of the return on such positions is driven by variation in jump premia.
2. Although far from perfect, the results of estimating the restricted model on option portfolio returns offers better results than estimating solely on option prices. This is due to the fact that the estimation objective function is tied directly to the dynamic properties of the returns.
3. Under correct specification the return-based approach provides very good estimates of the dynamic evolution of Sharpe ratios.

The evidence in figure 2 is pointing strongly to the fact that the type of data that can be accomodated in the estimation is of equivalent importance as the correctness of model specification and the choice of the objective function. In summary, all three should be tailored to the task at hand.

[3]The Sharpe ratios are calculated for an investment strategy that buys options with a given maturity (larger than 1 week), holds them for 1 week while delta-hedging the positions, and liquidates, taking in the proceeds from the sale. The option portfolios replicate the log contract.

2.2 Related literature

Our paper relates to the option implied risk premia literature as well as to the parametric model estimation literature. Starting with unspanned risk premia, Carr and Wu [2009] note that the VRPs on equity indices are predominantly negative and that there exists in the market a systematic variance risk factor whose price of risk is negative and which is unrelated to the three Fama and French [1993] factors. Looking at equation (1) it is straightforward to start thinking about the term structure of VRP by varying $T-t$. In this line of research Aït-Sahalia et al. [2012] found that the term structure slopes downwards most of the time, and that at the short end it's dominated by the component stemming from abrupt, jump-like movements. Andersen et al. [2015a] use the affine framework to determine that part of the VRP is driven by variables that do not drive the stock returns, and that shocks to VRPs are more persistent than shocks to asset price volatility. Bollersev et al. [2009] report that the VRP has some predictive power for excess stock returns, albeit in most studies the effect prevails only in-sample. This predictive power improves if jump and non-jump components of the variance risk premium are isolated (Li and Zinna [2014]). Finally Bardgett et al. [2013] estimate a model for two derivatives markets: the S&P 500 options and the VIX options. While their study reports that the affine models can't reproduce all the features of the markets, they also conclude – which is important in the light of our approach – that "the information contained in S&P 500 derivatives does not span the information contained in VIX derivatives and vice-versa". Our premise is similar, albeit it refers to two quantities observable in a single market: *the information contained in derivatives prices does not span the information contained in derivatives returns and vice-versa.*

Within the no-arbitrage framework the estimated risk premia translate directly into conditional expectations and variances of option portfolio returns, which allows us to answer questions about optimal investment with options and its relation to recent findings on variance swap trading strategies. Literature in this area is rather scarce. To the best of our knowledge, only Bakshi and Kapadia [2003] considered hedged returns on option positions to obtain an estimate of unspanned risk compensation, but they did not consider optimal trading strategies. Egloff et al. [2010] and Aït-Sahalia et al. [2012] provide evidence for variance swap positions – essentially delta-hedged option portfolios as well – and determine investors should sell short-term variance swaps and buy long-term variance swaps to obtain high Sharpe ratios in the $1.3 - 1.7$ range (annualised).

Turning to estimation, it can be shown that if both the $\mathbb{P}$ and $\mathbb{Q}$ measures allow an *affine representation* (see Duffie et al. [2000]), then the VRP is an *affine function* of the underlying variance factors. Thus, the affine stochastic volatility model, introduced in its general form in Duffie et al. [2000], is the most important tool in the analysis of variance risk premia and latent risk factors. Despite the relative analytical convenience of models in this class, their estimation is technically challenging and computationally expensive, especially if one considers higher-dimensional specifications. The nature of the problem is two-fold. First, one has to estimate the structural model parameters and second, simultaneously, one has to recover the values of latent factors.

In the literature we can discern a few approaches to estimating affine models which differ in the way they attempt tackle the latter problem, and in the type of data they employ (option prices or stock returns and variability measures based on high-frequency records). Pan [2002] and Andersen et al. [2015b], and Andersen et al. [2015a] use a GMM framework where they treat state variable values as additional parameters and obtain them by inverting *option prices* given a set of $\mathbb{Q}$-measure model parameters. As such, the resulting state space variable dynamics is in no way constrained, albeit in Andersen et al. [2015a] the resulting inverted estimates are shrunk towards estimates of spot variance obtained from high-frequency underling data. Other papers which concentrate on estimation methods do not necessarily pursue evidence on risk premia. Maximum likelihood estimation on stock return data has been considered by **?** and Bates [2006], who use stock return data and the analytical tractability of the characteristic function in affine models to solve the filtering problem. Rockinger and Semenova [2005] avoids the filtering problem by calculating the unconditional expectation of the characteristic function of stock returns. The aforementioned Aït-Sahalia et al. [2012] uses Kálmán filtering in a maximum likelihood framework and fitting to variance swap data, similarly to Egloff et al. [2010]. Non-linear filtering approaches have also been considered, Bardgett et al. [2013] is a state-of-the-art example.

We also note that our paper is also closely related to the pricing kernel projection literature. In particular, the unconditional pricing kernel projected onto stock-returns was introduced in Rosenberg and Engle [2002] which found the puzzling "U-shape" effect. Recent efforts by Song and Xiu [2012] have shown that conditioning the pricing kernel on variables such as the VIX index can resolve the shape puzzle. Our paper is related to this strand of literature in the sense that it uses a parametric model to produce the conditional pricing kernel. However, in our case the pricing kernel is defined as a function of unspanned risks, including multiple volatility factors. Also, our conditioning information is much richer, including (through the Kalman filter described in Section 4) the full history of observed returns as well as non-parametric measures of volatility.

Finally, the class of contracts which we introduce in section 3 is the same (up to a scaling factor) as the class of divergence portfolios introduced by ?.

3 Traded portfolios with analytical moments

The continuous-time option pricing literature distinguishes two sources of linearly unspanned risks: multifactor stochastic variance and jumps. These types of unspanned risks are naturally linked to different types of option portfolios (i.e. they can be traded by holding such portfolios and dynamically hedging in the underlying). For example, log contracts of varying maturities can be used to trade variance factors with different persistence. Log contracts also load on jump risk. However, they are not flexible enough to accurately trade the higher order risks implied by the distribution of return jumps. Portfolios which hold a larger number of put options than a log contract offer more exposure to negative jumps. Specialized trading strategies are required in order to preserve accurate delta-hedging properties. The analytical difficulties in treating individual options make them less tractable than they would seem.

We create exposure to different sources of risk by varying option portfolio composition in a way which renders the strategies easy to implement and delta hedge. The averages, variances and Sharpe ratios of returns on these strategies testify of the prices of the risks reflected in their payoffs.

Furthermore, some of these quantities are tractable in an affine stochastic volatility framework.

3.1 Power portfolios

To gain flexibility in engineering substantially different risk exposures while retaining analytical convenience we specify a class of delta-hedged *power portfolios*.

Definition 1 (power portfolio)**.** *A power portfolio with time to maturity $T-t$ and* frequency $u \in \mathbb{R}$ *is a static portfolio of out-of-the money options maturing at time T that replicates the payoff*

$$g\left(F_{T,T} \middle| u, F_{t,T}\right) := \exp\left(u(\log F_{T,T} - \log F_{t,T})\right) = \left(\frac{F_{T,T}}{F_{t,T}}\right)^u , \tag{5}$$

where $F_{t,T}$ is the futures price of the underlying at time t with time to maturity $T-t$.

The forward price of a power portfolio is the risk-neutral expectation of its payoff, $\mathbb{E}^{\mathbb{Q}}\left[g(F_{T,T}|u, F_{t,T})\right]$, which coincides with the time-t conditional Laplace transform of $F_{T,T}$ evaluated at u. An important consequence of definition 1 is that we can concisely describe the dynamics of the portfolio value together with its delta-hedging coefficients.

Corollary 1.

1. The futures value of the power portfolio struck at $\log F_{t,T}$ at time $s \geq t$ is given by:

$$\begin{aligned} P_{s,T-s}(u|\log F_{t,T}) &:= \mathbb{E}_s^{\mathbb{Q}}[g(F_{T,T}|u, F_{t,T})] \\ &= \mathbb{E}_s^{\mathbb{Q}}[\exp(u(\log F_{T,T} - \log F_{s,T}))] \cdot \exp(u(\log F_{s,T} - \log F_{t,T})) \\ &= \Psi^{\mathbb{Q}}_{s,T-s}(u) \cdot \exp(u(\log F_{s,T} - \log F_{t,T})) , \end{aligned} \tag{6}$$

   where $\Psi^{\mathbb{Q}}_{s,h}(u)$ is the risk-neutral conditional Laplace transform of the log futures return at time s over horizon $T-s$.

2. Under the assumption that option prices are homogenous in the stock price, the delta-hedging coefficient of the power portfolio is:

$$\frac{\partial P_{s,T-s}(u|\log F_{t,T})}{\partial F_{s,T}} = u \cdot \Psi^{\mathbb{Q}}_{s,T-s}(u) \cdot \frac{\exp\left(u(\log F_{s,T} - \log F_{t,T})\right)}{F_{s,T}} . \tag{7}$$

By immediate application of corollary 1, the futures prices itself is a power portfolio: $F_{t,T} = P_{t,T-t}(1|0)$.

3.2 Exposure to risk

The key feature of the power portfolios is that they load differently on put and call options, depending on the frequency u. From Carr et al. [2002], we know that the weights of the replicating OTM options as a function of the strike are given as:

$$w(K) = \frac{\partial^2}{\partial K^2} g(K) = u(u-1) F_{t,T}^{-2} \left(\frac{K}{F_{t,T}} \right)^{u-2} . \tag{8}$$

Depending on the frequency u the weight $w(K)$ is either tilted towards put or call options, generating different exposures to positive versus negative jumps.

To gain intuition on how variation in frequency (u) and maturity ($T-t$) affect the exposure of delta-hedged returns turn to Ito's lemma[4]:

$$P_{t+h,T-t-h}(u) - P_{t,T-t}(u) - \int_t^{t+h} \frac{\partial P_{\tau-,T-\tau}(u)}{\partial F_{\tau-,T}} dF_{\tau,T} = \frac{1}{2} \int_t^{t+h} \frac{\partial^2 P_{\tau-,T-\tau}(u)}{\partial F^2_{\tau-,T}} d[F_{\cdot,T}, F_{\cdot,T}]^c_\tau +$$
$$\sum_{i=1}^{N_j} \left(\Delta P_{\tau_i,T-\tau_i} - \frac{\partial P_{\tau_i-,T-\tau_i}}{\partial F_{\tau_i-,T}} \Delta F_{\tau_i,T} \right) + \int_t^{t+h} \frac{\partial P_{\tau-,T-\tau}(u)}{\partial V_{\tau-}} dV_\tau +$$
$$\frac{1}{2} \int_t^{t+h} \frac{\partial^2 P_{\tau-,T-\tau}(u)}{\partial^2 V_{\tau-}} d[V,V]_\tau + \frac{1}{2} \int_t^{t+h} \frac{\partial^2 P_{\tau-,T-\tau}(u)}{\partial F_{\tau-,T} \partial V_{\tau-}} d[F_{\cdot,T}, V]_\tau ,$$

where N_j is the number of jumps in the interval $(t, t+h)$ and τ_i, $i \in \{1, \ldots, N_j\}$ are the jump times. The second term on the right-hand-side captures the effect of return jumps. From Equation (6) we obtain[5]:

$$\Delta P_{\tau_i} - \frac{\partial P_{\tau_i-}}{\partial F_{\tau_i-}} \Delta F_{\tau_i} = \Delta P_{\tau_i} - u \frac{P_{\tau_i-}}{F_{\tau_i-}} \Delta F_{\tau_i} = \Psi^{\mathbb{Q}}_{\tau_i-}(u) \left(\frac{F_{\tau_i-,T}}{F_{t,T}} \right)^u \cdot [(1+r_{\tau_i})^u - 1 - u r_{\tau_i}] \tag{9}$$

where $r_{\tau_i} = \frac{F_{\tau_i}}{F_{\tau_i-}} - 1$. If we Taylor-expand in $r = 0$, we have up to the third

[4]For brevity, we assume here no jumps in variance and finite activity jumps in the underlying, and that there is a single variance factor.

[5]We drop maturity indexing for clarity of notation.

order:

$$(10)\ \ \Delta P_{\tau_i} - \frac{\partial P_{\tau_i -}}{\partial F_{\tau_i -}} \Delta F_{\tau_i} \approx \Psi^{\mathbb{Q}}_{\tau_i}(u) \left(\frac{F_{\tau_i -, T}}{F_{t,T}} \right)^u u(u-1) \left(\frac{1}{2} r^2_{\tau_i} + \frac{1}{6}(u-2) r^3_{\tau_i} \right)$$

This shows that the influence of jump-driven skewness is dependent on the frequency. For instance, $u = 2$ implies a vanishing dependence on jump=driven skewness. We note that the $u = 2$ portfolio corresponds to the Simple Variance Swap by Martin [2012].

Therefore by investigating delta-hedged returns on option portfolios with different frequencies, we can create different exposure styles to jumps, which can then by used to identify the parameters of the jump distribution relevant for unspanned risk.

Looking at the third term, we see the direct effects of stochastic multifactor volatility. From Equation (6) we see that the effect of stochastically changing volatility can only come from its effect on the $\mathbb{Q}$-conditional Laplace-transform. Therefore volatility factors with high persistence under the pricing measure will have a significant effect on long-maturity portfolios, whereas volatility factors with low persistence will only have minor effects. Therefore as in Egloff et al. [2010], considering different maturity portfolios can create exposures to different variance factors, assuming that they have different levels of persistence.

We note that we are not the first ones to entertain the investigation of power portfolios. Such a payoff has been considered already to be informative regarding the forward looking volatility dynamics, in particular Carr and Lee [2008] show that under an independence assumption, such power payoffs can be used as building blocks to replicated arbitrary realized volatility/variance payoffs. power portfolios are also a subclass of a broader family of power portfolios, whose properties are described by Schneider and Trojani [2015]

3.3 Model structure

We have seen that power portfolios are a flexible family that offer exposures to various unspanned risks. Now we also show that their first and second moment properties are easily calculable under affine option-pricing models. To fix ideas, we describe the assumed dynamics of the stock-price process and the latent factors under both the physical ($\mathbb{P}$) and risk-netural ($\mathbb{Q}$)

measures. To this end, let S_t denote the stock-price process and let $X_t = (V_t^1, ..., V_t^N)$ denote the N-dimensional latent Markov process of volatility factors. The assumed dynamics under the measure $\mathbb{M} \in \{\mathbb{P}, \mathbb{Q}\}$ is:

$$d \log S_t = \mu_S^{\mathbb{M}}(X_t)dt + \Sigma_S(X_t)dB_t^{S,X} + q_S^{\mathbb{M}} dN_S^{\mathbb{M}}(\lambda_S^{\mathbb{M}}(X_t)) \tag{11}$$

$$dX_t = \mu_X^{\mathbb{M}}(X_t)dt + \Sigma_X(X_t)dB_t^{S,X} + q_X^{\mathbb{M}} dN_X^{\mathbb{M}}(\lambda_X^{\mathbb{M}}(X_t)) \tag{12}$$

Where we have used the following notation:

- $B_t^{S,X}$ is an $N+1$ dimensional Brownian-motion driving stock and volatility innovations
- $\Sigma_X \in \mathbb{R}^{N \times N}$ and $\Sigma_S \in \mathbb{R}^N$ determine the instantaneous covariance of the stock and factor innovations. Under the assumptions of equivalence of the measures $\mathbb{P}$ and $\mathbb{Q}$ these quantities are the same under both measures.
- The quantities μ_S and μ_X determine the instantaneous drift of the stock and factor processes. Under the assumption of no-arbitrage,

  $$\mu_S^{\mathbb{Q}} = r_t - d_t - \frac{1}{2}\Sigma_S^T \Sigma_S - \lambda_S^{\mathbb{Q}} \mathbb{E}^{\mathbb{Q}} \left[\exp(q_S) - 1\right]$$

  where r_t is the instantaneous risk-free return and q_t is the continously compounded dividend yield at time t.
- The processes N_S and N_X are Poisson process with instantaneous intensities $\lambda_\cdot^{\mathbb{M}}$, $\cdot \in \{S, X\}$. Conditional on there being a jump, $q_\cdot^{\mathbb{M}}$ determines the random jump size, which also depends on the measure $\mathbb{M}$.

The analytical tractability of such multivariate processes is greatly increased under the assumption of an *affine structure*. As defined by Duffie et al. [2000], we say that the $N+1$ dimensional process $(\log S_t, X_t)$ is an affine process under measure $\mathbb{M}$ if $\mu_S^{\mathbb{M}}(\cdot), \mu_X^{\mathbb{M}}(\cdot), \Sigma_{S,X}\Sigma_{S,X}^T(\cdot), \lambda_S^{\mu}(\cdot)$ are affine functions of the factors X_t, where $\Sigma_{S,X}^T = [\Sigma_S^T | \Sigma_X^T]$.

3.4 Moments of power portfolios in affine models

In our context, the essential feature of power portfolios is their analytical tractability in affine models. The following proposition is the foundation for obtaining the covariance structure of Δ-hedged portfolios in affine-models:

Proposition 1. *Let $t_0 < t_1 < t_2 < ... < t_n$ and $i \in \mathbb{N}, i < \infty$. Then* ***in affine models*** *the moments of the following type can be calculated by solving a set of recursive ODEs (involving both $\mathbb{P}$ and $\mathbb{Q}$ parameters):*

$$M_{t_0}\left[(t_i, T_i, u_i, A_i, p_i)_i\right] := \mathbb{E}^{\mathbb{P}}_{t_0}\left(\prod_i \left[P_{t_i,T_i-t_i}(u_i|A_i)\right]^{p_i}\right). \tag{13}$$

3.5 The trading strategy

Hereby we describe the exact trading strategy and how it is delta-hedged. Assume that at time t the investor creates a power portfolio with frequency u and maturity $T-t$, the portfolio being struck at $\log F_{t,T}$. We will from now on assume that whenever the investor enters a power portfolio position, it is struck at the current log futures price, for notational convenience. As discussed above, the futures price of the portfolio is $P_{t,T-t}(u|\log F_{t,T})$. Now if the investor holds this portfolio over horizon h and sells it on the market[6], then his *unhedged return* – ρ – is:

$$\rho_{t,h}(u) := \frac{P_{t+h,T-t-h}(u)}{P_{t,T-t}(u)} - 1 = \frac{\Psi^{\mathbb{Q}}_{t+h,T-t-h}(u)}{\Psi^{\mathbb{Q}}_{t,T-t}(u)} \cdot \left(\frac{F_{t+h,T}}{F_{t,T}}\right)^u - 1 \tag{14}$$

From (14) it is clear that besides taking on the risk of the state variables moving (which affect Ψ), the investor also has substantial exposure to the movement in the futures price. To avoid this, we assume the investor enters into Δ-hedging on the futures market of the stock. The power portfolio is easy to delta hedge exactly. We remind here corollary 7, which states that the instantaneous sensitivity to the current futures price is[7]:

$$\frac{\partial P_{\tau,T}(u)}{\partial F_{\tau,T}} = \frac{u}{F_{\tau,T}} P_{\tau,T}(u). \tag{15}$$

Using the above observation, we can write the Δ-hedged returns – r – as

[6]Note that we assume *no trading* in the options over the interval $[t, t+h]$, i.e. the investor holds exactly the same portfolio at the end as at initiation.

[7]Since $\Psi_{t,T-t}(u)$ depends only on the volatility state factors and not on the underlying (futures) level.

the following (assuming a hedging frequency of δ):
(16)

$$r_{t,h,\delta}(u) \equiv \frac{P_{t+h,T-h}(u) - u\sum_{j=0}^{h/\delta} \frac{P_{t+j\delta,T-h-j\delta}(u)}{F_{t+j\delta,T}} \left(F_{t+(j+1)\delta,T} - F_{t+j\delta,T}\right)}{P_{t,T-t}(u)} - 1\,.$$

Hence, we investigate the behaviour of investors who at time t purchase power portfolios with maturity $T-t$ and hold them for time h, then liquidate. Investors delta hedge the portfolios by trading in the stock futures market every δ units of time.

Note that all of the terms in (16) are of the form described in Proposition 1, therefore evaluating the unconditional or conditional $\mathbb{P}$ expectation is straightforward given the model parameters. If one considers unconditional or conditional second moments, then expanding all the terms, they are still a linear combination of elements of the form in Proposition 1. Even though the number of terms in the expansion of r^2 grows at a rate of δ^{-2}, it is possible to calculate an accurate approximation at a low computational cost by applying sparse quadrature approximations as detailed in the Appendix.

While the unconditional expectations of the moments of $r_{t,h}$ are the easiest to calculate, it is not the return specification that is the most economically and economterically appealing. In particular if one takes the Carr-Madan expansion of the power-portfolio weights, then it turns out that the replicating portfolio contains a unit amount of bond besides the pure out-of-the money option portfolio. Whilst this constant bond amount does not change the delta-hedged payoff, it does change the delta-hedged return. From an economic perspective it is more intuitive therefore to consider the following quantity, which is similar to the logarithmic return on variance swaps introduced in Carr and Wu [2009]:

$$\begin{aligned} y_{t,h}(u) &= \frac{P_{t+h,T-h}(u) - u\sum_{j=0}^{h/\delta} \frac{P_{t+j\delta,T-h-j\delta}(u)}{F_{t+j\delta,T}} \left(F_{t+(j+1)\delta,T} - F_{t+j\delta,T}\right) - P_{t,T-t}(u)}{P_{t,T-t}(u) - 1} \\ &= r_{t,h}(u)\frac{P_{t,T-t}(u)}{P_{t,T-t}(u) - 1} \end{aligned}$$

We also found that from an econometric noise-specification point of view (see Section 4 on the Kálmán filter) this specification was better suited. In particular, it is more reasonable to propose that the error standard deviations coming from imperfect option prices are constant across time for y than for r. To provide an analogy, it is more reasonable that errors in the VIX squared imputed from option prices are proportional to the VIX squared level, than saying that they are independent from it and are constant across time.

Naturally both the model implied conditional moments and the data implied returns are easily calculable for y as well[8]. Based on the above described numerical tractability of conditional first and second moments, we describe in the next section how to set up a Kálmán filter to efficiently infer state dynamics and construct a quasi-likelihood criterion for parameter optimization.

4 Return based estimation

The guiding philosophy of latent state models is that there is a small set of unobservable factors that solely drive observed variables. In option pricing, these states are usually associated with stochastic variance.

An option pricing model that specifies both $\mathbb{P}$ and $\mathbb{Q}$ dynamics plays two roles when estimating unspanned risks. First, it allows to *calculate estimates of the latent states* (the filtering step). Secondly it allows to estimate risk prices *given the estimated states* (the forecast step). This gives rise to multiple viewpoints on how a model should be estimated. We can differentiate between these along two dimensions.

The first distinction is between the "maximalists" and the "minimalists". The "maximalist" would posit that given a full description of both $\mathbb{P}$ and $\mathbb{Q}$ measures, one should use all available underlying and derivative information to estimate a model efficiently, which implies the usage of non-linear filtering methods to estimate states. The "minimalist" contends that such requirements are too stringent both from a computational and a specificational point of view. Model misspecification is inevitable and thus the model should be tailored to the task at hand, trading off efficiency in other areas. Using all available data increases model complexity and computational cost. The "minimalist" thus considers a model that's a "good enough" description of the phenomenon under study.

[8]The "leverage" quantity $\frac{P_{t,T-t}(u)}{P_{t,T-t}(u)-1}$ is known at time t, hence we don't calculate its expectation.

The second distinction is between a “descriptive” and a “forecasting” model. The “descriptive” approach estimates a model that posits the joint dynamics of the observables, assuming that the model is correctly specified, and then uses the a filtering technique to build the likelihood. Alternatively, a simple filtering device – such as the Kálmán filter – provides a “regularized” set of state (i.e. their conditional first and second moments). The filter is treated as a “forecasting” device which is judged on the basis of its empirical practicality rather than its statistical soundness. The hereby defined “forecasting” approach selects the model parameters which provide the best in-sample conditional second moment forecasts for variables of interest.

4.1 Statement of the estimation objective

Let $m_{t,t+h}$ denote the conditional pricing kernel between time t and $t+h$. It is defined as the random variable which obeys the conditional moment equality:

$$\mathbb{E}_t^{\mathbb{P}}\left[m_{t,t+h} r^e_{t,t+h}\right] = 0\,, \tag{17}$$

where $r^e_{t,t+h}$ is *any* excess return over $[t, t+h]$, specifically a return on a derivatives position. It is accepted practice to assume that the pricing kernel is a linear function of realized returns:

$$m_{t,t+h} = a\left(V_t\right) + \mathbf{b}\left(V_t\right) \cdot \mathbf{r}_{t,t+h}\,, \tag{18}$$

where $a \in \mathbb{R}$ and $\mathbf{b} \in \mathbb{R}^{1\times n}$ are functions of the state vector V_t. Our goal is to accurately estimate the dynamic features of a and $\mathbf{b}$, which requires (i) a reasonable approximation of the functions a and $\mathbf{b}$ and (ii) a plausible description of the dynamic properties of V_t.

The assumption of a linear pricing kernel, though restrictive at first sight, is one we are willing to make. It is the simplest formulation which can correctly price $n+1$ observed market returns, so to set the bar higher we extend the set of returns that have to be priced correctly. The literature to-date has not considered a dynamic estimation of such a pricing kernel for a set of returns on derivative positions as described in section 3.

The functions a and $\mathbf{b}$ are determined by the conditional first and second moments of the returns. Thus, our approximate pricing kernel retains all the information that is relevant for a myopic mean-variance investor. As concisely summarised by Cochrane [2014], mean-variance analysis "is a good robust first step, even if the second step is fairly large as well." Therefore, our objective in the estimation is to target the conditional second moment structure of returns on derivative positions.

4.2 Kálmán filter

We are able to describe the joint conditional second moment properties of the returns on the underlying asset, returns on power portfolios, their prices (optionally, not in all estimations), and the variance factors (and functions thereof, e.g. bipower variation). The Kálmán filter is a natural tool in such a setting. Under admittedly stringent conditions it is the optimal predictor under quadratic loss, but even if – as in this paper – not all of these conditions are met, it remains the best *linear* predictor. The use of the filter allows us to impose intertemporal constraints on the dynamics of the factors, so that the whole available time series of the data up to time t is used for prediction of latent states at time t.

The breadth of our information sets puts us halfway between the "maximalist" and "minimalist" approaches: on the one hand we use high frequency data on index returns in delta hedging and bipower variation, on the other we refrain from fitting individual option prices, resorting to option portfolio prices, if necessary. As evident from estimation results in section 6, portfolio prices – summary statistics of option prices – impose inflexible restrictions on factor values. While beneficial in a well-specified setting, such a feature is not necessarily favorable as it reduces the model's ability to fit the dynamic properties of *returns*. In consequence of these design choices, our model is decisively of the "forecasting" guise.

4.2.1 Observed quantities

We use the following quantities in estimation: (i) *subsequent weekly returns* on the S&P 500 index; (ii) *subsequent weekly returns* on U different continuously delta-hedged option portfolios; (iii) start-of-the week prices of the U portfolios; (iv) time series of bipower variation (BV), observed at the end of each corresponding trading period.

Bipower variation is essential as a regularisation device along the lines of Andersen et al. [2015b] if portfolio prices are excluded from the observation set. BV is a linear function of the variance states with coefficients that are model parameters themselves, and not functions thereof; it also is independent of the model's pricing assumptions, as long as the model precludes arbitrage. Enriching the dataset with S&P 500 index returns facilitates the identification of return jump features.

Let $r_{t,h} = (r^0_{t,h}, y^1_{t,h}, ..., y^U_{t,h})$ denote the return on the underlying and U distinct delta-hedged power portfolios. We assume that the returns $r_{t,h}$ are observable at time $t+h$ and observed every $h = 5/252$ years, i.e. every week, on Wednesdays. This choise is driven by holiday considerations (there are fewest on Wednesdays) as well as spread sizes in the options market. It is not conceivable to trade options at much higher frequencies than weekly without incurring very high transaction costs. Finally, option prices themselves contain considerable noise, therefore returns over shorter horizons would be dominated by the noise effects.

The option position are hedged over the course of the week as in equation 16 during the opening hours of the futures market at a 5-minute frequency. The overnight hedge results from the last futures position of the day being held until the market reopens. The hedge ratios are established with the latest available option data, i.e. mostly the previous day's closing option prices, except for the overnight hedge, for which the power portfolio value can be recalculated. In this way delta hedging introduces information contained in high-frequency data into the estimation process. Furthermore, unlike previous studies, we do not include only one functional of the high frequency data (such as realized variance), but we include U different transforms, corresponding to U different option portfolios. Therefore our filter incorporates more information from stock returns than most other methods considered before.

4.2.2 Latent states

We assume that there are two different sources of uncertainty influencing the option and underlying returns. The first group of effects comprises the changes in the $N + 1$ dimensional vector of stochastic factors $(\log S_t, V_t)$. In other words, returns are driven by changes is in the variance factors and the underlying.

Secondly, we assume that option prices are *observed with errors.* The observed *errors in the prices* manifest in *autocorrelated errors in the returns*. This autocorrelation needs to be taken into account in the observation equations discussed below. For reasons that will become apparent below, we treat this group of errors are latent states.

Let V_t denote the time-t value of the demeaned variance factors and let $\epsilon_t^R = (\epsilon_t^i)_{i \in \{1..U\}}$, denote the scaled error in the prices of the power portfolios, while ϵ_t^P denotes portoflio price observation errors themselves. [9]. The price errors are assumed to be IID normal, with constant covariance matrix Ξ^R. We fill in details on the error specification in section 5, however we note that in general we can expect cross-correlation in observation errors for portfolios that are calculated from options with the same maturity. This is due to the fact that in the replicating portfolio we have the same options, albeight with different weights for different frequencies.

The variance factors are assumed to be governed by the following propagation equation:

$$\Delta V_{t+h} = \mu_V(V_t) + \Sigma_V(V_t)\epsilon_{t+h}^V \tag{19}$$

where $\epsilon_{t+h}^V \sim N(0, h \cdot \mathbb{I}) \perp \epsilon_{t+h}^R$. The functions μ_V and Σ_V are easily calculable in affine models. In practive, when implementing the filter, we will input the previously filtered value $\hat{V}_t$ into $\mu_V(\cdot)$ $\Sigma_V(\cdot)$ when evaluating the propagation equation.

[9]Here we are agnostic about the exact nature of the pricing errors. In particular, we will define the errors based on their effect on the hedged returns. However we call them price errors to motivate the autocorrelation effect they cause in the hedged returns.

4.2.3 Observation equations

The observation equations have the following form:

$$
\begin{aligned}
(20)\qquad r_{t,h} &= \mu_R(V_t, \epsilon_t^R) + H(V_t)\Delta V_{t+h} + \epsilon_{t+h}^O + \epsilon_{t+h}^R \\
&= \mu_R(V_t) + H(V_t)\Delta V_{t+h} + \epsilon_{t+h}^O + \left(\epsilon_{t+h}^R + \rho\epsilon_t^R\right) \\
(21)\qquad \log P_{t,t+h} &= \alpha_P + \beta_P V_t + \epsilon_t^P \\
(22)\qquad \mathrm{BV}_{t+h} &= \sum_{i=1}^{N} V_{t+h}^i + \epsilon_{t+h}^{BV} \\
\text{with } \mu_R\left(V_t\right) &= \mathbb{E}_t\left[r_{t,h}\right] , \\
H\left(V_t\right) &= \mathbb{V}_t\left[\Delta V_{t,t+h}\right]^{-1} \mathrm{cov}_t\left[r_{t,h}, \Delta V_{t,t+h}\right] .
\end{aligned}
$$

The matrix H and vector μ_R defined as the L_2 projection[10] of the returns on the innovations in the latent factors V_{t+h}. These are easily calculated via proposition 1 in affine models. We note that the returns depend on the variance shocks, whilst BV depends on the *variance levels*. This is important for to the specification of the filtered covariance matrix of the levels of the states, which do not enter into the conditional covariance matrix of the returns, but do enter into the conditional covariance matrix of BV.

Accordingly to the L_2 projection, the shock ϵ_{t+h}^O – which represents the uncertainty in returns conditional on a given outcome of the factors – is assumed to be uncorrelated with $\epsilon_{V,t+h}$ (and also with ϵ_{t+h}^R). ϵ_{t+h}^O is the error of the L_2 projection with variance $\mathbb{V}\left[r_{t,h}\right] - H(V_t)\ \mathrm{cov}_t\left[r_{t,h}, \Delta V_{t,t+h}\right] H(V_t)^T$. The final term in the observation equations is the moving average term coming from the previous period's pricing error and the current period's pricing error. The auto-correlation coefficient matrix $\rho \in \mathbb{R}^{(U+1)\times(U+1)}$ is treated as a parameter in later estimation. The error in portfolio prices ϵ_t^P is, for convenience, modeled as an IID process correlated with $\epsilon_{t+h}^R + \rho\epsilon_t^R$.

At this point it is also good to draw the attention to an additional way our filter differs from that of Egloff et al. [2010]. In their paper, the observed *error structure* of the realized variance conditional on the factor values was estimated seperately of the parameters. Contrary to their approach, in our case the $\mathbb{P}$ parameters also specify what the conditional variance of the realized variance should be. In other words, the covariance matrix of ϵ_{t+h}^O is dependent on V_t and the model parameters and is not assumed to be

[10]Imagine calculating the betas of returns with respect to the factors.

constant and is not estimated separately from the parameters.

4.2.4 Likelihood function

Given the state propagation and observation equations, we can build the likelihood function based on the conditional Gaussian approximation of the innovations.

To this end let $(\bar{x}_{t,h}, \bar{Q}_{t,h})$ denote the conditional mean and conditional covariance of bi-power variation and the $U+1$ returns (recall that the returns are over the interval $[t, t+h)$) in our observation set, produced by the filter based on data up to time t. Then the conditional log-likelihood of observing the outcome $x_{t,h}$ at time $t+h$ is:

$$(23)\quad l_{t+h}(x_{t,h}|\Theta^{\mathbb{P}}, \Theta^{\mathbb{Q}}) = -\frac{1}{2}\left[\log|\bar{Q}_{t,h}| + \left((x_{t,h} - \bar{r}_{t,h})^T \bar{Q}_{t,h}^{-1}(x_{t,h} - \bar{r}_{t,h})\right)\right]$$

The total log-likelihood of the time series of returns can be calculated as the sum of the conditional log-likelihoods. The parameters $\Theta = (\Theta^{\mathbb{P}}, \Theta^{\mathbb{Q}})$ can then be estimated by numerically maximizing the total log-likelihood function.

We note that this is an approximate likelihood function – as in Egloff et al. [2010] – for two reasons. Firstly, in the Kálmán filter, the state propagation equation matrices should not depend on the previous state values. Secondly, the innovations ϵ^O, ϵ^V are generally not normally distributed.

There is a vast literature on the properties of the quasi-maximum likelihood estimator, starting from results on its consistency under certain assumptions[11] (Gourieroux et al. [1984], Bollerslev and Wooldridge [1992]) to results on its efficiency properties. In general the advantages of QMLE are its simplicity, computational speed, interpretability and potential robustness to misspecification. By the latter we mean that if the model flexible enough to mimick the second order features of the observations, then the estimated model implied second order quantities (e.g. Sharpe-ratios) will be consistent for the true ones even if the full likelihood specification of the model is incorrect. On the other hand, if one uses a true likelihood based estimation, then even the second order properties of the estimated model might be incorrect, given that the model is misspecified.

[11] These assumptions are generally not met in the case of square-root diffusions.

The empirical evidence on the approximate quasi-maximum likelihood setup[12] that we follow is generally supportive even if theoretically the estimates are potentially inconsistent. Duan and Simonato [1999] is an early example of this approach and they show that in their setting of modeling interest rates, the approximate QMLE on the CIR model has very similar features to a true MLE on the Vasicek model. They also verify that the finite sample properties (they consider samples of size 150-400 months) of the estimator are reasonably described by the asymptotic approximations. Egloff et al. [2010] also report obtaining good fitting variance swap rates and reasonable model implications using the approximate QMLE method. Finally, the linearized approximate Kálmán filter is recommended by Duffee and Stanton [2012] as a reasonably efficient alternative to MLE (with similar small-sample bias properties) and they suggest that it can even be superior in inference accuracy.

Given that our goal in this paper is to investigate the second order properites of trading various option based portfolios and the high probability of fitting mis-specified models[13], it is natural that we use QMLE to estimate our models. In order to demonstrate the properties of the approach in our setting, we will perform an extensive Monte Carlo analysis in Section 5. Nevertheless, it remains an interesting open question to see how "sharpening the likelihood" (for an example of using importance sampling to move away from QMLE to MLE, see Brandt and He [2006]) can help in better pinpointing the higher order properties of returns.

5 Monte Carlo exercise

We illustrate the properties of our estimation method via a Monte Carlo study. We are in particular interested in seeing how well our approach is able to capture the dynamics of risk premia and the covariance structure of option returns, especially in a mis-specified model. We set up a base three factor model and estimate it both under the correct specification and under a constrained one, where one of the non-jumping factors is removed. Then

[12] As discussed, our estimation approach is only an approximate QMLE, in the sense that the conditional moments of the states and observations are dependent on the latent states.

[13] To date there is no model to the best of our knowledge that has been convincingly proven to be flexible enough to describe the full $\mathbb{P}$ and $\mathbb{Q}$ properties of the S&P 500 market over the full observation horizons available to researchers.

we proceed to comparing our estimation method against a competing technique: QML-type estimation with state filtering from option prices serves as a good backdrop for highlighting the strengths and weaknesses of our approach. The concepts of central interest in this exercise are the recovery of conditional variance risk premia and of the latent state variables, which determine Brownian-type and jump risk.

5.1 Three factor SV model

As the data generating process we use a three-factor model whose specification is inspired by the recent findings of Andersen et al. [2015a] and Carr and Wu [2011], who first and foremost conclude that the performance of option pricing models greatly improves with the introduction of a third driving factor. The model allows for simultaneous asset price and variance jumps, and its most prominent feature is that the stochastic jump intensity process is partially decoupled from the continuous stochastic variance of the asset price.

We lay down the model structure under the pricing measure $\mathbb{Q}$, and follow with parameter restrictions and risk premia specifications. In the most general case we could consider, asset price and variance factor dynamics follow the stochastic differential equations:

$$
\begin{aligned}
\frac{dS_t}{S_{t-}} &= \mu_t^{\mathbb{Q}} dt + \phi_1 \sqrt{V_{1t}} dB_{1t}^{\mathbb{Q}} + \phi_2 \sqrt{V_{2t}} dB_{2t}^{\mathbb{Q}} + \phi_3 \sqrt{V_{3t}} dB_{3t}^{\mathbb{Q}} + \left(e^{J_S} - 1\right) dN_t \\
dV_{1t} &= \kappa_1^Q \left(\eta_1^{\mathbb{Q}} - V_{1t}\right) dt + \sigma_1 \sqrt{V_{1t}} dW_{2t}^{\mathbb{Q}} + J_{V_1} dN_t \\
dV_{2t} &= \kappa_2^Q \left(\eta_2^{\mathbb{Q}} - V_{2t}\right) dt + \sigma_2 \sqrt{V_{2t}} dW_{2t}^{\mathbb{Q}} \\
dV_{3t} &= \kappa_3^Q \left(\eta_1^{\mathbb{Q}} - V_{3t}\right) dt + \sigma_3 \sqrt{V_{3t}} dW_{3t}^{\mathbb{Q}}
\end{aligned}
$$

(24)

$$\mu_t^{\mathbb{Q}} = r - q + \alpha \left(\sum_{n=1}^{3} \phi_n \sqrt{1 - \rho_n^2} V_{nt} \right) - \lambda_t \left(\theta_J^{\mathbb{Q}}(e_1) - 1 \right) ,$$

with jump intensity

$$\lambda_t = \lambda_J + \lambda_{J,1} V_{1t} + \lambda_{J,2} V_{2t} + \lambda_{J,3} V_{3t} ,$$

jump measure

$$\theta(c) = \frac{\mu_{J_S} c_1 + 1/2\, \sigma_{J_S}^2 c_1^2}{1 - \mu_{J_V} c_2 - \rho_{J_V} \mu_{J_V} c_1}, \text{ and}$$

$$\operatorname{cov}(dB_i, dW_j) = \rho_j, \quad \operatorname{cov}(dW_1, dW_2) = 0, \quad \operatorname{cov}(dB_2, dB_3) = 0$$

with prices of risk

$$\Gamma_t^i = \frac{\gamma_1^i}{\sqrt{V_i}} + \gamma_2^i \sqrt{V_i} \, .$$

This general structure has far too many parameters to believe that they can be identified and consistently estimated on a sample of about 500 weekly observations. We impose restrictions on risk premia specifications and selected parameters of model (24) in order to improve parameter identifiability and reproduce important features observed in empirical data:

1. $\phi_1 = 0$: the jumping variance factor does not directly drive the stock price variance;
2. $\lambda_{J,3} = 0$: the third factor does not influence the jump intensity;
3. $\gamma_1^2 = \gamma_2^2 = 0$: [14] the second factor is purely statistical, unpriced risk, hence $\kappa_2^{\mathbb{P}} = \kappa_2^{\mathbb{Q}}$ and $\eta_2^{\mathbb{P}} = \eta_2^{\mathbb{Q}} = 1$;
4. $\gamma_1^3 = 0$: the third factor mean-reversion rate is priced, while it's long-term means isn't, and hence $\eta_3^{\mathbb{P}} = \eta_3^{\mathbb{Q}} = 1$;

[14] The γ_j^i parameters are defined implicitly by choosing κ and η parameters under both measures.

5. γ_1^1 is such that $\eta_1^{\mathbb{P}} = 1$ for identification purposes;

6. λ parameters are identical under $\mathbb{Q}$ and $\mathbb{P}$ measures[15].

Partial variance / intensity decoupling via (1.) and (2.) above stems from observation by Andersen et al. [2015a] that the pricing of jump risk is not uniquely determined by the Brownian part of asset price variance. The restrictions in (3.) reflect their finding that there exists a fast-moving variance factor, that does drive risk, but doesn't drive its compensation. Via (4.) and (6.) we reduce the number of estimated parameters by 5.

We choose parameters so that the jump-driving factor V_1 exhibits a half-life of approximately 7 weeks and that abrupt increases in variance cause an even more persistent change in risk premia (18-week mean reversion under the pricing measure). The "pure risk" factor V_2 has a half-life of 3 weeks and high volatility-of-volatility. The "mid-term risk" factor V_3 gets a half-life of 36 weeks under $\mathbb{P}$ and 50 weeks under $\mathbb{Q}$. All parameters are given in table 1 below.

When estimating a two-factor constrained specification, we allow all γ_j^i risk premia parameters to be non-zero, so that the model remains as flexible as possible, in order to capture most variance risk premium variation patterns.

5.2 Generation of hedged returns

We simulated model (24) for 1000 weeks at a 5-minute frequency to obtain the time series of variance factors and the high-frequency futures observations. Given these "observations" and the model parameters, we created option panels. To best replicate the features of the S&P 500 options market we used data available in the OptionMetrics database. We regressed every day's and every maturity's time- and volatility corrected maximum and minimum available log-strike prices $\left(\frac{k}{\sigma_{IV}\sqrt{\tau}}\right)$ on polynomials in σ_{IV} and τ. Thus we learned what ranges of options are actively quoted and traded at

[15] As jumps are rarely observed, the $\mathbb{P}$ parameters of the jump intensity are very hard to pin down if they differ from the $\mathbb{Q}$ parameters, this way their identification is strengthened by the fact that for a given intensity level that matches scant $\mathbb{P}$ evidence, high option prices will first impact the $\mathbb{Q}$ jump parameters and then contribute to the intensity identification. Finally, due to the fact that the jump intensity depends on the volatility processes, $\mathbb{E}_t^{\mathbb{P}}[\lambda_{t+s}] \neq \mathbb{E}_t^{\mathbb{Q}}[\lambda_{t+s}]$, i.e. the jump intensity also carries a risk premium that is induced by the risk premia on variance processes.

	Parameter	Value	Parameter	Value
1	$\alpha_1^{\mathbb{P}}$	0.0000	$\eta_1^{\mathbb{Q}}$	1.0250
2	$\eta_1^{\mathbb{P}}$	1.0000	$\kappa_1^{\mathbb{Q}}$	2.0000
3	$\kappa_1^{\mathbb{P}}$	5.0000	$\sigma_1^{\mathbb{Q}}$	2.0000
4	$\sigma_1^{\mathbb{P}}$	2.0000	$\phi_1^{\mathbb{Q}}$	0.0000
5	$\phi_1^{\mathbb{P}}$	0.0000	$\rho_1^{\mathbb{Q}}$	-0.7000
6	$\rho_1^{\mathbb{P}}$	-0.7000	$\eta_2^{\mathbb{Q}}$	1.0000
7	$\eta_2^{\mathbb{P}}$	1.0000	$\kappa_2^{\mathbb{Q}}$	12.0000
8	$\kappa_2^{\mathbb{P}}$	12.0000	$\sigma_2^{\mathbb{Q}}$	4.8500
9	$\sigma_2^{\mathbb{P}}$	4.8500	$\phi_2^{\mathbb{Q}}$	0.1500
10	$\phi_2^{\mathbb{P}}$	0.1500	$\rho_2^{\mathbb{Q}}$	-0.9000
11	$\rho_2^{\mathbb{P}}$	-0.9000	$\eta_3^{\mathbb{Q}}$	1.0000
12	$\eta_3^{\mathbb{P}}$	1.0000	$\kappa_3^{\mathbb{Q}}$	0.7180
13	$\kappa_3^{\mathbb{P}}$	1.0000	$\sigma_3^{\mathbb{Q}}$	1.4100
14	$\sigma_3^{\mathbb{P}}$	1.4100	$\phi_3^{\mathbb{Q}}$	0.1200
15	$\phi_3^{\mathbb{P}}$	0.1200	$\rho_3^{\mathbb{Q}}$	-0.4000
16	$\rho_3^{\mathbb{P}}$	-0.4000	$\lambda_{J,1}^{\mathbb{Q}}$	2.0000
17	$\lambda_{J,1}^{\mathbb{P}}$	2.0000	$\lambda_{J,2}^{\mathbb{Q}}$	0.3333
18	$\lambda_{J,2}^{\mathbb{P}}$	0.3333	$\lambda_{J,3}^{\mathbb{Q}}$	0.0000
19	$\lambda_{J,3}^{\mathbb{P}}$	0.0000	$\lambda_J^{\mathbb{Q}}$	0.1000
20	$\lambda_J^{\mathbb{P}}$	0.1000	$\nu_J^{\mathbb{Q}}$	0.2500
21	$\nu_J^{\mathbb{P}}$	0.1500	$\mu_J^{\mathbb{Q}}$	-0.0500
22	$\mu_J^{\mathbb{P}}$	-0.0200	$\rho_J^{\mathbb{Q}}$	-0.1000
23	$\rho_J^{\mathbb{P}}$	-0.0500	$\sigma_J^{\mathbb{Q}}$	0.1000
24	$\sigma_J^{\mathbb{P}}$	0.0400		

Table 1: Simulation model parameters

each maturity. We used the regression model to determine the strike ranges for simulated option prices given volatility factor values. We assumed that 50 options are observed at every maturity, which is an underestimate in most of the latter part of the sample, but an overestimate in the earlier part of the sample. Finally, we decided to use four maturities (1, 3, 6 and 12 months) in the simulated sample, whilst the true market offers richer data sets with 7 to 12 maturities usually available. The option prices were simulated within a fixed-expiry framework, where new options are issued every month once the maturity of the nearest option decreases to less than one week, so that our dataset contains some calendar effects in terms of data availability. An i.i.d. noise with standard deviation 15% in relative prices (which is generally more noise than observed in real data) was also added to all option observations to mimic the fact that market option observations also contain noise.

For the test estimation we used frequencies $u \in \{-1, 3\}$ and maturities $\tau \in \{1/12, 6/12, 1\}$. The choice of the negative frequency aims at identifying asset price jumps, while the choice of the positive frequency at pinning down how variance factor moves influence asset price skewness. Thus for every day in the sample we calculated the requisite option price transforms and interpolated their values to maturities τ.

Finally, we split the sample in half and used only 500 observations for estimation, while the other 500 were used to evaluate the model's portfolio optimization performance.

Taking the interpolated option transforms and the high-frequency data on the underlying stock we proceeded with calculating hedged returns along formula 16.

5.3 Estimation results

There are multiple dimensions to our comparison of the estimated models. First, within the scope of a single estimation method, we compare the filtering performance of a properly and an improperly specified model. We repeat this comparison in the scope of recovering the term structure of conditional variance risk premia. Then we take a different point of view: within the scope of a given (correct or incorrect) specification, we compare performance in the aforementioned tasks across estimation methods.

The parameters that maximized the total log-likelihood over our sample in the three factor model return-based estimation are summarized in Table 2. The return noise covariance matrix in the estimation was assumed to be block-diagonal (with diagonal values $\exp(\epsilon_i^P), i = 1 \ldots 6$), with contemporaneous correlations between portfolios with same maturities (ρ_1^C,ρ_2^C,ρ_3^C). The autocorrelation in the noise was assumed to be maturity dependent as well, implying three additional parameters (ρ_1^A,ρ_2^A,ρ_3^A).

In general we see that even though we only used returns on 6 different portfolios and estimated the models on samples of 500 observations, the estimated parameters are mostly inline with the true ones reported in Table

1. Tables with the other models' estimated parameters can be found in the Appendix.

	Parameter	Value	Parameter	Value
1	$\kappa_1^{\mathbb{P}}$	5.8152	$\eta_1^{\mathbb{Q}}$	0.8294
2	$\sigma_1^{\mathbb{P}}$	2.0340	$\kappa_3^{\mathbb{Q}}$	0.9025
3	$\rho_1^{\mathbb{P}}$	0.1326	$\nu_J^{\mathbb{Q}}$	0.1458
4	$\phi_2^{\mathbb{P}}$	0.1303	$\rho_J^{\mathbb{Q}}$	0.0376
5	$\kappa_2^{\mathbb{P}}$	12.2666	$\mu_J^{\mathbb{Q}}$	-0.1279
6	$\sigma_2^{\mathbb{P}}$	5.5171	$\sigma_J^{\mathbb{Q}}$	0.0630
7	$\rho_2^{\mathbb{P}}$	-0.8995	ρ_1^A	-0.2896
8	$\phi_3^{\mathbb{P}}$	0.1472	ρ_2^A	-0.7744
9	$\kappa_3^{\mathbb{P}}$	1.9171	ρ_3^A	-0.6728
10	$\sigma_3^{\mathbb{P}}$	1.3204	ρ_1^C	0.0978
11	$\rho_3^{\mathbb{P}}$	-0.3588	ρ_2^C	0.8511
12	$\lambda_J^{\mathbb{P}}$	0.2786	ρ_3^C	0.5560
13	$\lambda_{J,1}^{\mathbb{P}}$	1.9514	ϵ^{BV}	-6.7604
14	$\lambda_{J,2}^{\mathbb{P}}$	0.2633	ϵ_1^P	-9.0524
15	$\nu_J^{\mathbb{P}}$	0.0452	ϵ_2^P	-9.0229
16	$\rho_J^{\mathbb{P}}$	-0.1765	ϵ_3^P	-9.9461
17	$\mu_J^{\mathbb{P}}$	-0.0083	ϵ_4^P	-8.8574
18	$\sigma_J^{\mathbb{P}}$	0.0368	ϵ_5^P	-8.9583
19	$\kappa_1^{\mathbb{Q}}$	1.7337	ϵ_6^P	-9.0731

Table 2: Estimated parameters using the return-based calibration on the Monte Carlo sample of 500 weeks, using the correctly specified model.

5.4 Estimating risk premia

Empirical evidence suggests that the variance risk premium is negative and has an approximately flat unconditional term structure. In periods of low aggregate volatility the term structure can become negatively sloped, and it also has been observed that after severe market disruptions the slope can be positive Gruber et al. [2015].

Figures 3 through 7 illustrate how well the models fare in matching market risk premia. First, in the top-left corner of figure 3 we notice that the option-return estimated three-factor model matches the unconditional variance risk premium closest, and that the two-factor option-return-estimated model performs almost as well as the richer three-factor specification, if estimated on option prices. Again, the model exhibits good performance in capturing the part of the jump variance risk premium due to a pure-jump measure difference (i.e. $\mathbb{E}^{\mathbb{P}}\left[\int_t^T \lambda_s ds\right]\left(\mathbb{E}^{\mathbb{P}}\left[J_S^2\right] - \mathbb{E}^{\mathbb{Q}}\left[J_S^2\right]\right)$), but in the pure intensity part (i.e. $\left(\mathbb{E}^{\mathbb{P}}\left[\int_t^T \lambda_s ds\right] - \mathbb{E}^{\mathbb{Q}}\left[\int_t^T \lambda_s ds\right]\right)\mathbb{E}^{\mathbb{P}}\left[J_S^2\right]$) the the results favor other approaches.

The dynamic properties of the term structure of variance risk premia are again better replicated by the option-return-estimated models. The left panels of figures 4 through 7 show the alignment of true and estimated values of the level and slope of the variance risk premium. While the price-estimated three factor model performs best in this respect, it is worth noting that the two-factor return-estimated model paints a better picture than the same model estimated on option prices. In terms of matching the comovement in VRP slope and level (right panels of figures 4 through 7), the return-based models have a clear advantage, especially when comparing the two-factor, misspecified ones. Both price-estimated models exhibit less variation in the VRP term structure level given a VRP term structure slope, and the estimates from the two-factor model based on prices are significantly biased.

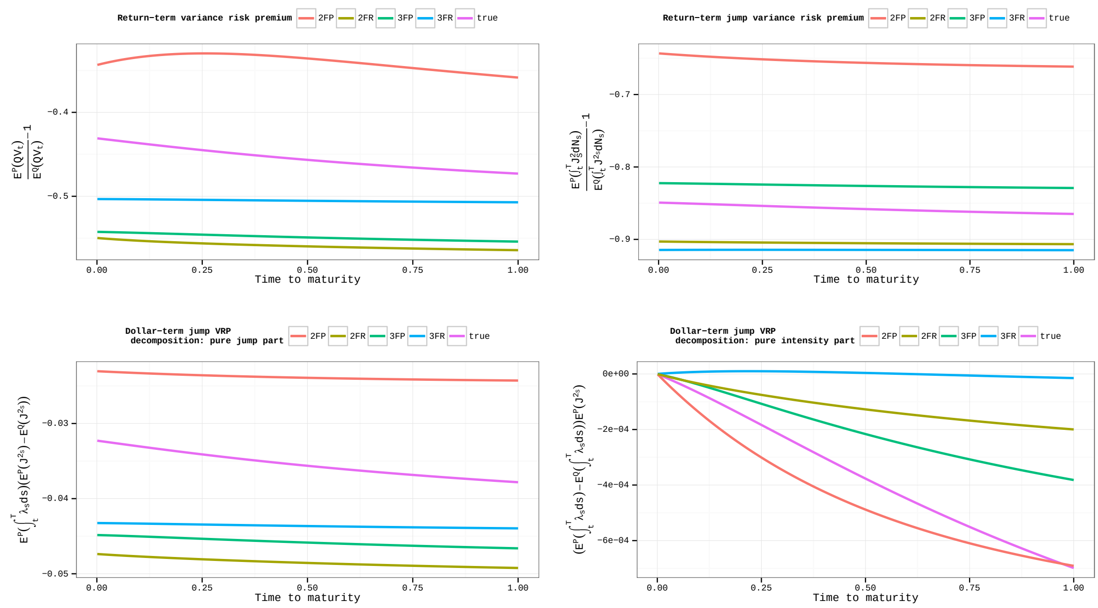


Figure 3: Unconditional term structures of variance risk premia and decompositions across models.

5.5 Filtering variance factors

Recovering latent variance factors accurately is critical for assessing current market risks, forecasting them and calculating option portfolio sensitivities. In figure 8 we present true and estimated factor values from the three-factor model. The jump factor (V_1), which doesn't drive the Brownian part of the stock dynamics, is estimated well, but some imperfections in its filtering enter the estimates of the mid-term risks factor (V_3), which, however, has a smaller influence on current asset price variation than the pure risk factor (V_2), which is estimated very precisely. Ultimately, the total conditional continuous variance estimate is close to the truth. The model does a similarly commendable job when recovering the conditional jump intensity, which loads most strongly on the jump factor.

In comparison, as presented in figure 9, in the two-factor estimation on option returns, the estimates of the conditional continuous variance are very close to the true values, while there is a clear downward bias in the estimates of the conditional jump intensity – the model clearly lacks the flexibility to match the high level of jump variation and a relatively lower level of continuous variation in the original model. Both the "filtered factors" exhibit reasonable dynamics and they differ strongly in mean-reversion speeds. The first filtered factor loads mostly on the non-jumping true factors (V_2 and V_3), and is negatively correlated with the jumping factor, while the second, slower mean-reverting factor is estimated as the primary driver of the jumps, and loads on V_1 and V_3.

Estimating the three-factor specification on option prices with the use of the Kálmán filter yields results that are not significantly different from option-return-based estimation. A few qualitative differences are worth noting, however. In the second panel of figure 10 it is clearly visible that the filtering fails to capture some of the variation in the pure risk factor V_2, while the jumping factor is recovered as good as in the option-return-based estimation, except for periods when it's very high. This bias arises most probably because in the option-returns-based filtering the observation equations are closer to being linear than in the option-price-based filtering. Finally, it is worth noting (in the fifth panel of figure 10) that estimating a correctly-specified model on option prices yielded very precise estimates of the conditional jump intensity.

The good results of option-price-based estimation do not transfer to a mis-specified model. The results of the filtering exercise are presented in figure 11. The model's ability to recover conditional Brownian variance is significantly hampered, as evident in the third panel of figure 11. At the same time the conditional jump intensity is significantly overestimated across the whole sample.

This part of the MC exercise clearly indicates that while for a properly specified model option-price based estimation yields similar results to option-return-based estimation, the latter method is highly superior in improperly specified settings. The good performance of the three-factor option-price-estimated model indicates that the breakdown that we observe in the case of the two-factor specification should not be attributed to the fact that we use the Kálmán filter to recover the states.

All in all we conclude from the simulation that the proposed return based estimation method produces parameter estimates, dynamic risk premium forecasts and trading signals that are close to their true values and which are economically meaningful. This is even true for the case where the model is potentially misspecified. In contrast, traditional price based estimation fails to produce good results for the dynamic evolution of risk premia, in the case when the sample does not come from the assumed model.

6 Empirical results

Among others, Bakshi et al. [2003] and Egloff et al. [2010] document that there exists a term structure of risk premia associated with holding option portfolios. Bollerslev and Todorov [2011] suggest that jumps are the dominant driver of these premia. In order to investigate the expected returns and Sharpe ratios available for SPX option traders, we estimated the model on a 400-week subsample of a data set of 749 weeks of trading.

We employ the model specification from section 5.1 with all parameter constraints therein.

6.1 The dataset

The dataset contains option prices (extracted from the OptionMetrics database) and corresponding high-frequency price records of futures on the S&P 500 index. Our records span the period from September 1997 to February 2012, whilst for the presented estimations we use subsamples starting in January 2001 and November 1998 and ending in August 2008, before the Financial Crisis. The samples contain, respectively, 400 and 570 weekly option return observations. The option panels were sampled every Wednesday or, if unavailable, on the closest trading day. We applied standard filters to option prices and corrected them for non-convexity in a minimal manner. Interest rates and dividend yields were interpolated from the OptionMetrics data.

The weekly returns were calculated for frequencies $u \in \{-1, 3\}$ and maturities $t \in \{1/12, 1/2\}$. We found that calculating the prices of power portfolios for frequencies below -1 is highly sensitive to the smoothing method.

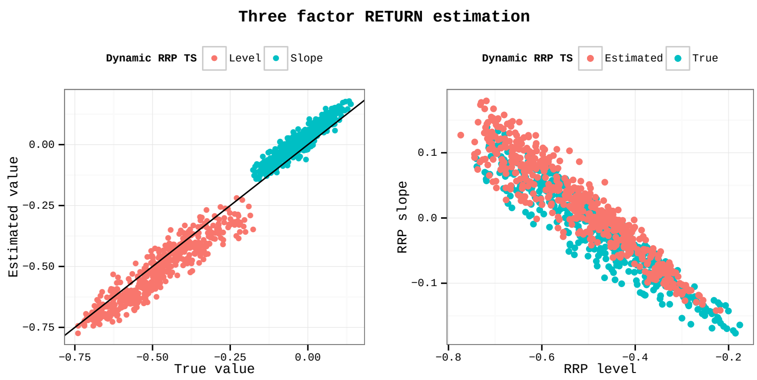


Figure 4: Slope $\left(RRP_{t,t+1} - RRP_{t,t+1/12}\right)$ and level $\left(RRP_{t,t+1/12}\right)$ of the return variance risk premium in simulated data and filtered from the return-estimated three factor model.

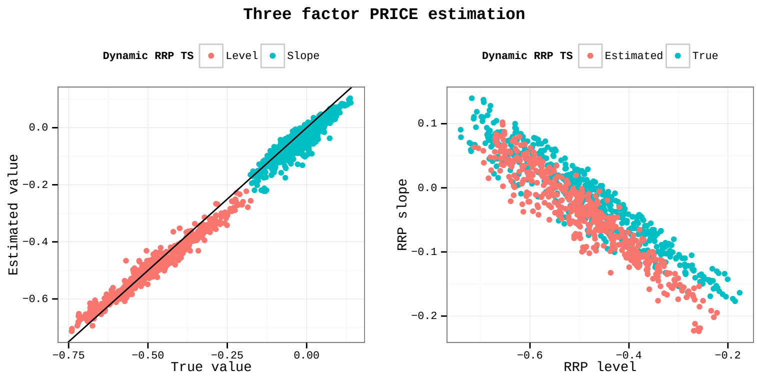


Figure 5: Slope $\left(RRP_{t,t+1} - RRP_{t,t+1/12}\right)$ and level $\left(RRP_{t,t+1/12}\right)$ of the return variance risk premium in simulated data and filtered from the price-estimated three factor model.

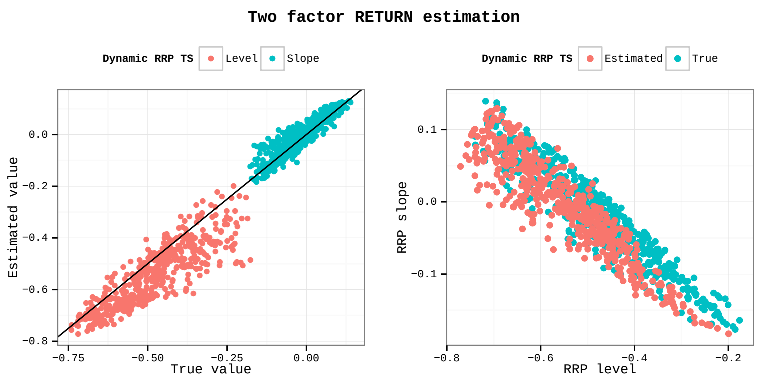


Figure 6: Slope $\left(RRP_{t,t+1} - RRP_{t,t+1/12}\right)$ and level $\left(RRP_{t,t+1/12}\right)$ of the return variance risk premium in simulated data and filtered from the return-estimated two factor model.

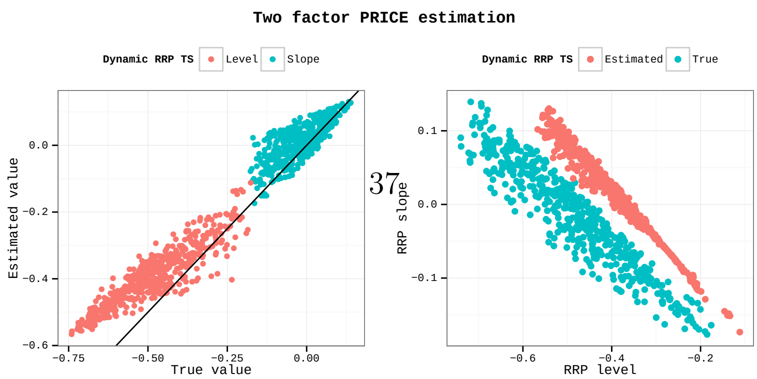


Figure 7: Slope $\left(RRP_{t,t+1} - RRP_{t,t+1/12}\right)$ and level $\left(RRP_{t,t+1/12}\right)$ of the return variance risk premium in simulated data and filtered from the price-estimated two factor model.

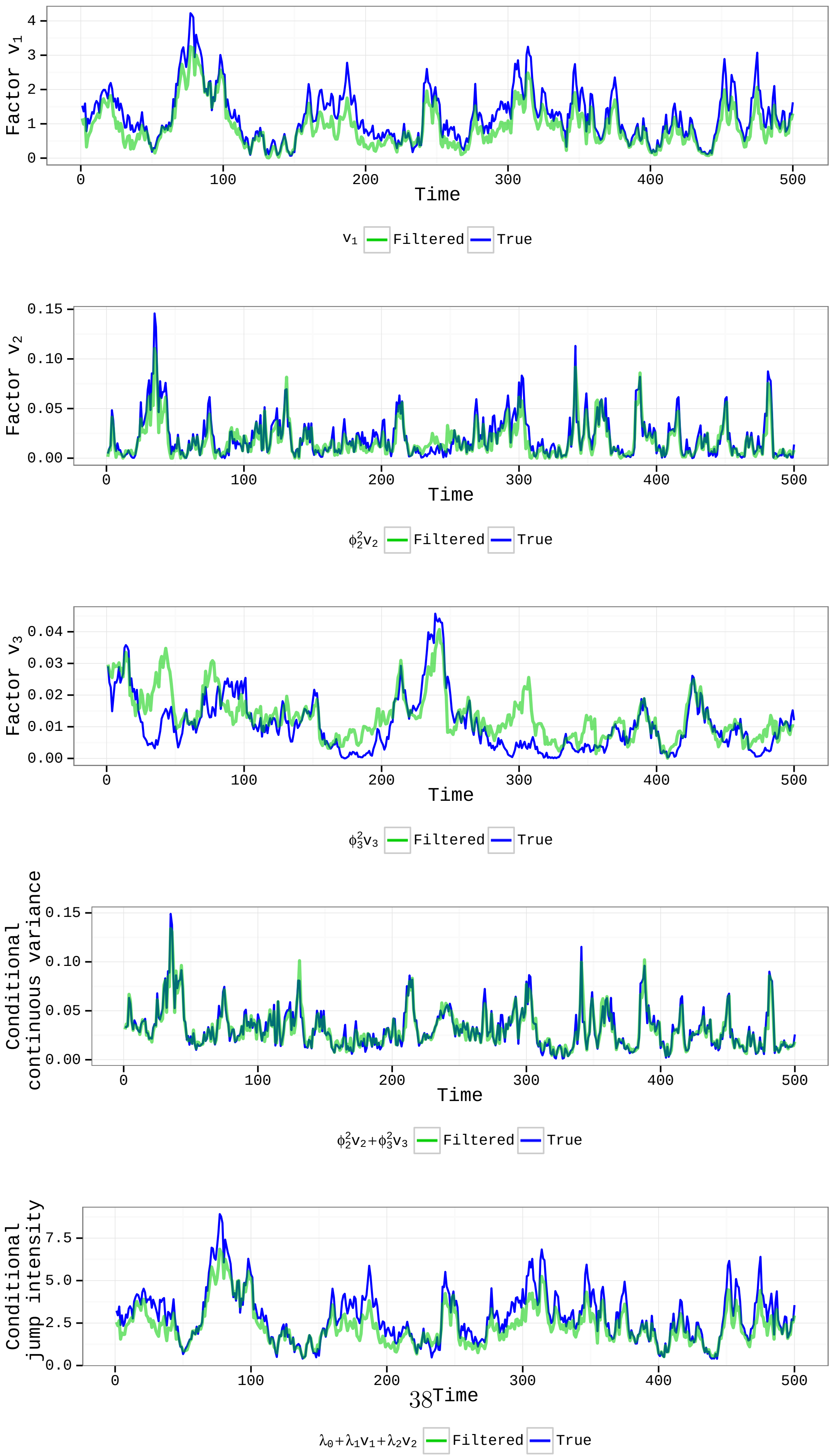


Figure 8: Simulated and filtered factors, conditional Brownian variance and conditional jump intensity, estimated in a three-factor model on option portfolio returns.

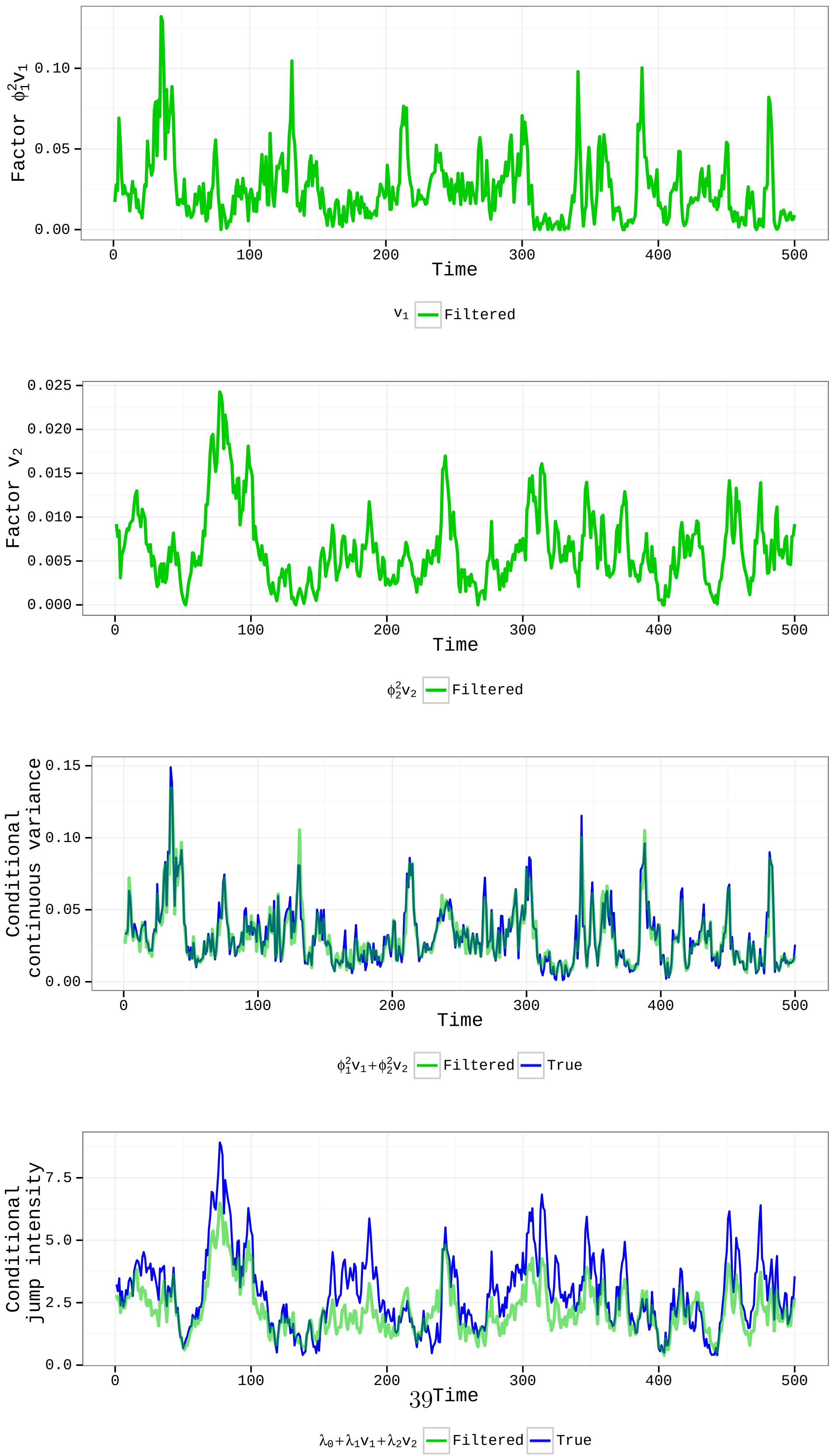


Figure 9: Filtered factors, conditional Brownian variance (filtered and true) and conditional jump intensity (filtered and true), estimated in a *two-factor* model on option portfolio returns.

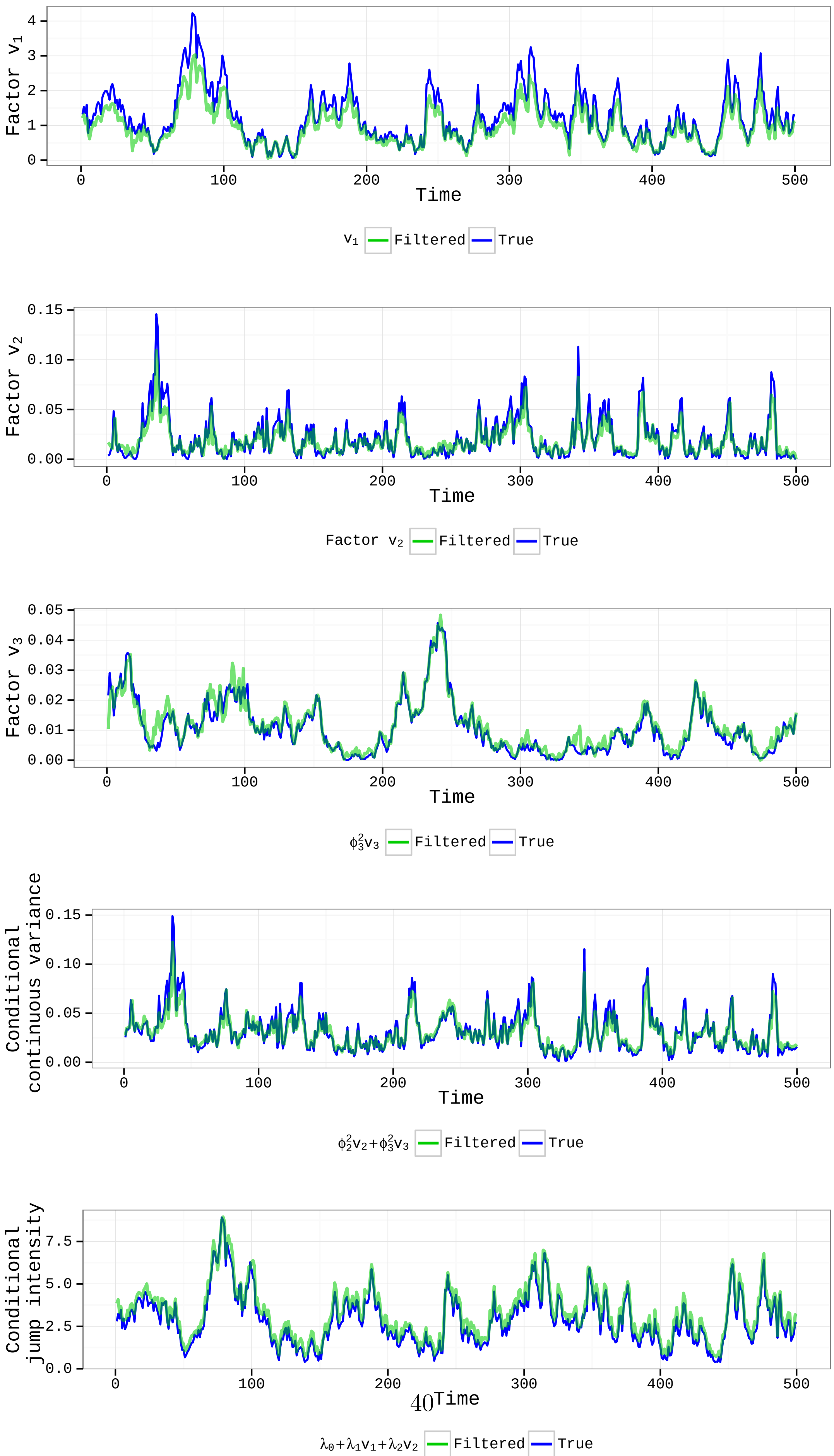


Figure 10: Simulated and filtered factors, conditional Brownian variance and conditional jump intensity, estimated in a three-factor model on option prices.

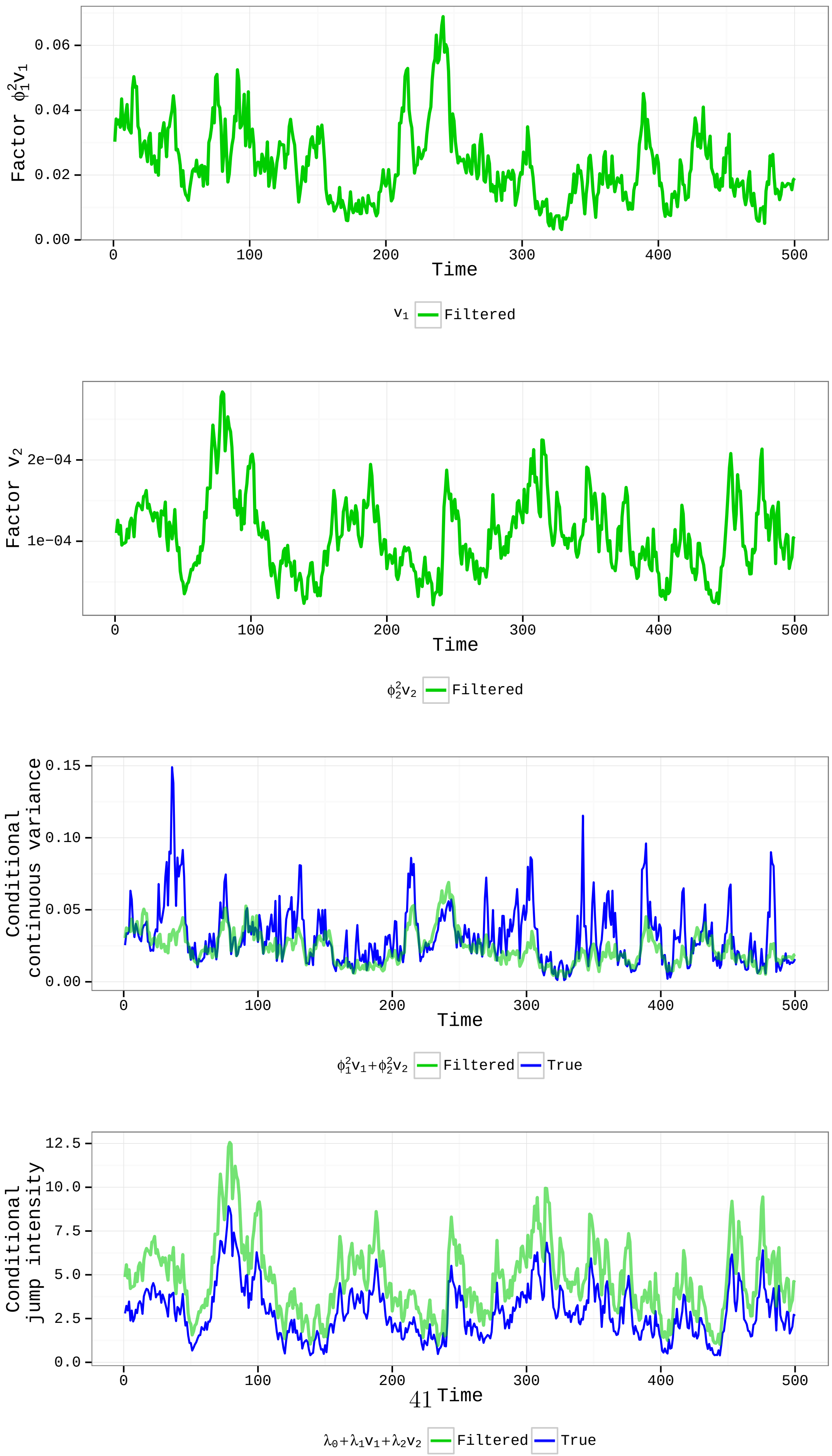


Figure 11: Filtered factors, conditional Brownian variance (filtered and true) and conditional jump intensity (filtered and true), estimated in a *two-factor* model on option prices.

	Returns					Leveraged returns				
Sample	**Statistic**	$r_{1M,-1}$	$r_{1M,3}$	$r_{6M,-1}$	$r_{6M,3}$	**Statistic**	$r_{1M,-1}$	$r_{1M,3}$	$r_{6M,-1}$	$r_{6M,3}$
1997-2012	$\mathbb{E}[r]$	−0.000381	−0.000761	−0.000604	−0.000862	$\mathbb{E}[r]$	−0.0747	−0.0612	−0.00823	−0.00998
	$\mathbb{V}[r]$	0.00000607	0.0000241	0.000257	0.000119	$\mathbb{V}[r]$	0.114	0.0818	0.0353	0.0161
	SR	−0.155	−0.155	−0.0376	−0.079	SR	−0.221	−0.214	−0.0438	−0.0787
	Min.	−0.0234	−0.0387	−0.155	−0.0661	Min.	−0.618	−0.626	−0.465	−0.342
	$q_{0.05}$	−0.00289	−0.00722	−0.01	−0.0155	$q_{0.05}$	−0.374	−0.373	−0.199	−0.178
	Med.	−0.000361	−0.000748	−0.000859	−0.00122	Med.	−0.142	−0.113	−0.0381	−0.0271
	$q_{0.95}$	0.00175	0.00496	0.00855	0.0138	$q_{0.95}$	0.381	0.382	0.243	0.19
	Max.	0.0233	0.0437	0.2	0.0693	Max.	5.13	3.14	1.85	1.06
	γ	1.98	1.28	1.89	0.892	γ	6.49	3.44	4.12	1.61
	κ	44.2	24.9	62.9	10.7	κ	80.7	26.1	32.1	8.24
1997-2008	$\mathbb{E}[r]$	−0.000352	−0.000767	−0.000488	−0.000807	$\mathbb{E}[r]$	−0.0801	−0.0664	−0.0108	−0.0106
	$\mathbb{V}[r]$	0.00000176	0.0000106	0.0000382	0.0000622	$\mathbb{V}[r]$	0.0756	0.0666	0.0235	0.0129
	SR	−0.266	−0.236	−0.0789	−0.102	SR	−0.291	−0.257	−0.0703	−0.0933
	Min.	−0.00677	−0.0166	−0.054	−0.0564	Min.	−0.561	−0.626	−0.45	−0.342
	$q_{0.05}$	−0.0022	−0.00545	−0.00686	−0.0123	$q_{0.05}$	−0.362	−0.365	−0.175	−0.168
	Med.	−0.000316	−0.000695	−0.000656	−0.00102	Med.	−0.135	−0.11	−0.0342	−0.0239
	$q_{0.95}$	0.00128	0.00363	0.00603	0.0112	$q_{0.95}$	0.319	0.336	0.202	0.186
	Max.	0.0101	0.0252	0.0714	0.0368	Max.	2.15	1.88	1.85	0.498
	γ	1.47	1.1	1.45	0.00287	γ	3.19	2.32	4.07	0.901
	κ	13.6	11.7	48.5	7.14	κ	17.8	10.5	39.6	2.06
2008-2012	$\mathbb{E}[r]$	−0.000476	−0.000744	−0.00098	−0.00104	$\mathbb{E}[r]$	−0.0574	−0.0444	0.00003	−0.00793
	$\mathbb{V}[r]$	0.0000201	0.0000682	0.000971	0.000304	$\mathbb{V}[r]$	0.24	0.131	0.0739	0.0264
	SR	−0.106	−0.0901	−0.0314	−0.0595	SR	−0.117	−0.123	0.00011	−0.0488
	Min.	−0.0234	−0.0387	−0.155	−0.0661	Min.	−0.618	−0.609	−0.465	−0.307
	$q_{0.05}$	−0.00502	−0.00982	−0.028	−0.0263	$q_{0.05}$	−0.418	−0.405	−0.264	−0.203
	Med.	−0.000629	−0.00126	−0.00199	−0.00252	Med.	−0.16	−0.121	−0.055	−0.0378
	$q_{0.95}$	0.00425	0.00882	0.0329	0.0278	$q_{0.95}$	0.418	0.423	0.389	0.262
	Max.	0.0233	0.0437	0.2	0.0693	Max.	5.13	3.14	1.84	1.06
	γ	1.34	0.921	1.1	0.954	γ	7.19	4.42	3.37	2.25
	κ	15.1	10.7	16.6	4.63	κ	70.5	33.3	16.9	10.8

Table 3: Summary statistics of returns and leveraged returns on delta-hedged option portfolios. γ denotes sample skewness, and κ denotes excess kurtosis.

Table 3 presents descriptive statistics of the returns. The returns are negative, positively skewed, and for a given frequency u increase in magnitude with maturity[16]. The pattern is more pronounced for leveraged returns, whose variance also decreases with maturity. Reported Sharpe ratios are for weekly holding periods and have not been annualized.

As visible in figure 1 on page 4, returns can take extreme values with kurtosis over 45 (returns) and 80 (leveraged returns) over the whole sample. In leveraged terms the extremes range from losing more than 60% on the investment or quintupling it *over 7 days*.

6.2 SPX market estimation results

6.2.1 Parameters

Table 4 presents parameters estimated on 400 weeks of trading data. The estimated model follows the observation structure of section 4.2.3, except for equation (21), which is removed from the specification.

We observe that the self-exciting jumping factor exhibits extreme persistence and high volatility. It's filtered values are mostly 0 during the conundrum period, and start exhibiting variation after 2005. It's the primary driver of jumps with 1.3 such movements per year expected in sample, but even up to 15 at the height of the financial crisis.

Mean variance jump size ν_J is small under $\mathbb{P}$, but important under $\mathbb{Q}$. The size of variance jumps drives the average size of asset price jumps, which are unconditionally distributed as $N(-0.0007, 0.052^2)$ under $\mathbb{P}$ and $N(-0.057, 0.073^2)$ under $\mathbb{Q}$.

Annualized diffusive volatility is approximately 15%, whereas total annualised volatility is 20.6%, very close to what Li and Zinna [2014] report.

The risk-neutral mean of the self-exciting jump intensity factor is 3.929, substantially higher than the $\mathbb{P}$ mean which is constrained to 1. Even though the factor mean reversion is faster under $\mathbb{Q}$ than under $\mathbb{P}$, it carries a significant part of the jump risk premium. The third, purely diffusive factor contributes to the variance risk premium at longer horizons via extremely slow $\mathbb{Q}$ mean reversion.

[16]In summary statistics not reported here for the 1 year options, the pattern is broken during the financial crisis.

The model reproduces the leverage effect prevalent in the data both in the diffusive and in the jump part. $\mathbb{Q}$ jump leverage is much stronger than $\mathbb{P}$ jump levereage, implying that risk-neutral skewness plays an important role in the pricing of power portfolios.

	Parameter	Value	Parameter	Value
1	$\alpha_1^{\mathbb{P}}$	0.019	$\lambda_{J,2}^{\mathbb{P}}$	0.002
2	$\kappa_1^{\mathbb{P}}$	0.012	$\lambda_J^{\mathbb{P}}$	0.000
3	$\sigma_1^{\mathbb{P}}$	4.049	$\nu_J^{\mathbb{P}}$	0.012
4	$\kappa_2^{\mathbb{P}}$	3.626	$\rho_J^{\mathbb{P}}$	-0.062
5	$\sigma_2^{\mathbb{P}}$	5.940	$\sigma_J^{\mathbb{P}}$	0.052
6	$\phi_2^{\mathbb{P}}$	0.113	$\eta_1^{\mathbb{Q}}$	3.929
7	$\rho_2^{\mathbb{P}}$	-0.986	$\kappa_1^{\mathbb{Q}}$	1.725
8	$\kappa_3^{\mathbb{P}}$	1.962	$\kappa_3^{\mathbb{Q}}$	0.036
9	$\sigma_3^{\mathbb{P}}$	0.987	$\nu_J^{\mathbb{Q}}$	0.297
10	$\phi_3^{\mathbb{P}}$	0.102	$\rho_J^{\mathbb{Q}}$	-0.192
11	$\rho_3^{\mathbb{P}}$	-0.809	$\sigma_J^{\mathbb{Q}}$	0.073
12	$\lambda_{J,1}^{\mathbb{P}}$	0.477		

Table 4: Estimated three factor model parameters, Θ_M

6.2.2 Expected option portfolio returns

The sample means of model-implied expected option returns, conditional variances and Sharpe ratios are close to their respective estimation sample counterparts[17], see table 5. The estimated model overestimates (in magnitude) expected returns on longer-maturity options and hence their Sharpe ratios. For short-maturity options, surprisingly, the model-implied average return on the $u = 3$ portfolio with 1 month maturity is estimated as higher than for $u = -1$. Table 6 contains model-implied unconditional quantities for the whole sample.

A visual investigation of model-implied option portfolio return Sharpe ratios (figure 12) reveals certain interesting patterns. For short-maturity option portfolios, they are almost perfectly correlated. Slight differences only become noticable during market turmoil. The trough in the Sharpe

[17]These aren't perfect counterparts: average conditional variance doesn't equal the conditional variance, similarly a sample average of conditional Sharpe ratios doesn't equal the Sharpe ratio constructed from unconditional sample mean and variance of returns

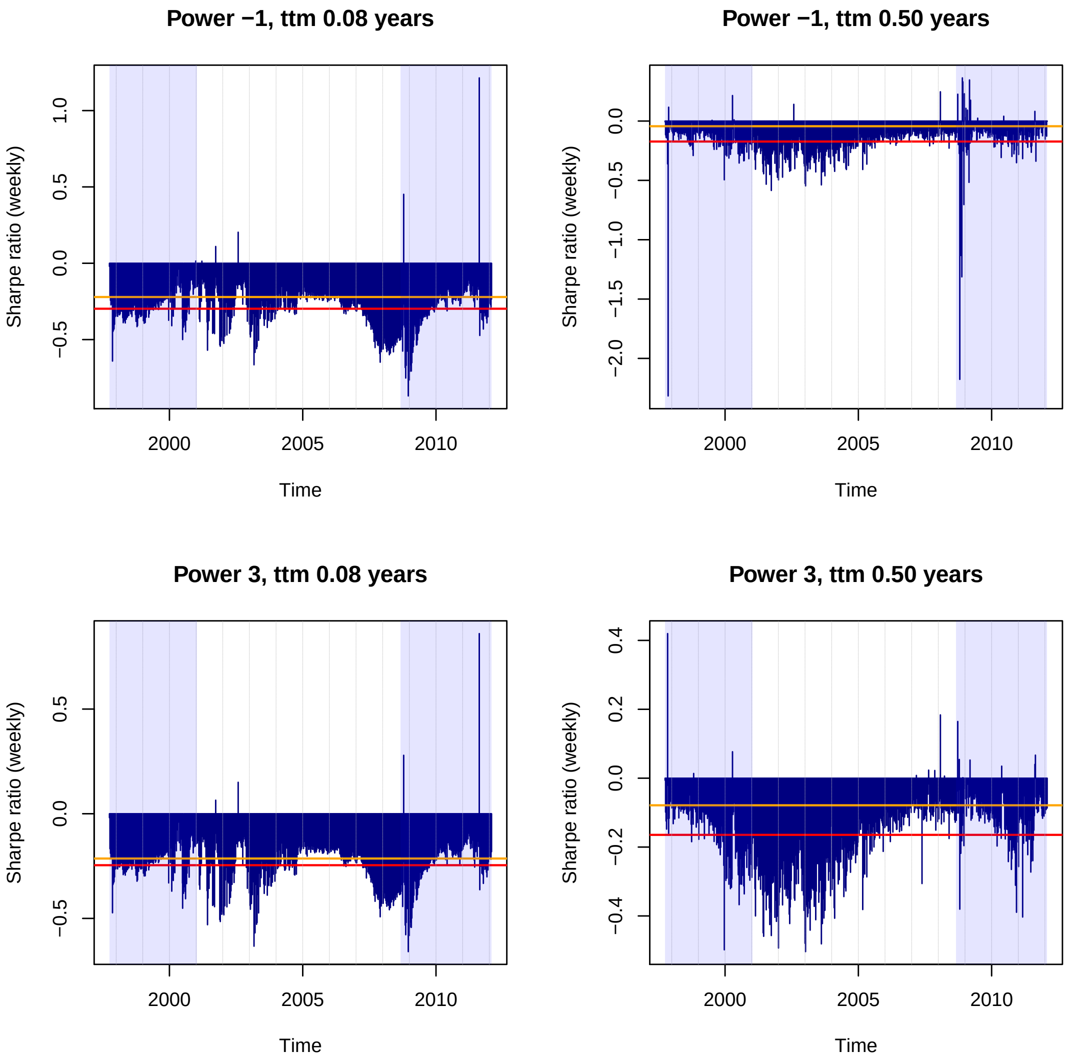


Figure 12: Model-implied option portfolio Sharpe ratios, blue shading denotes the out-of-sample period. Several positive predicted Sharpe ratios are consequences of very high variance factor values and consequently high returns.

	$\tau = 1M, u = -1$	$\tau = 1M, u = 3$	$\tau = 6M, u = -1$	$\tau = 6M, u = 3$
$\mathbb{E}\left[\mathbb{E}_t[y_{t,h}]\right]$	-0.0706	-0.0617	-0.0209	-0.0203
$\mathbb{E}\left[\mathbb{V}_t[y_{t,h}]\right]$	0.0691	0.0812	0.0138	0.0134
$\mathbb{E}[SR_t]$	-0.305	-0.258	-0.196	-0.198
$\mathbb{E}[y_{t,h}]$	-0.0747	-0.0593	-0.0126	-0.00974
$\mathbb{V}[y_{t,h}]$	0.0612	0.0568	0.0153	0.0125
$SR[y_{t,h}]$	-0.302	-0.249	-0.102	-0.0871

Table 5: Sample means of model-implied conditional expected option returns, conditional option return variances and conditional option return Sharpe ratios, estimation sample. Estimation sample return moments and SR added for convenience.

	$\tau = 1M, u = -1$	$\tau = 1M, u = 3$	$\tau = 6M, u = -1$	$\tau = 6M, u = 3$
$\mathbb{E}\left[\mathbb{E}_t[y_{t,h}]\right]$	-0.0742	-0.0645	-0.0196	-0.0181
$\mathbb{E}\left[\mathbb{V}_t[y_{t,h}]\right]$	0.0892	0.101	0.0243	0.0242
$\mathbb{E}[SR_t]$	-0.298	-0.246	-0.173	-0.165

Table 6: Sample means of model-implied conditional expected option returns, conditional option return variances and conditional option return Sharpe ratios, full sample.

ratios of $u = -1$ and 1-month maturity portfolios is slightly lower than for $u = 3$. For the longer maturity, 6 months, the differences are more pronounced, especially after extreme observations during the financial crisis.

The Sharpe ratios exhibit a time-varying term structure. It's flat during the conundrum period and upwards-sloping if diffusive-type risks dominate in the investor perception. During the financial crisis, especially for $u = -1$, the term structure is strongly downwards-sloping. No such effect is observed for $u = 3$.

The realized leveraged returns on power portfolios display extremely high kurtosis, which cannot be captured by the Kálmán filter in our estimation approach, especially when the variance factor values are high. As a consequence, the quality of conditional return prediction suffers. Table 7 offers a breakdown of correlation between predicted and realized returns by BV quartiles and subsample. The rightmost column of the table reports the overall correlation: interestingly, higher out-of-sample than in-sample. Most importantly the predicability of power portfolio returns is higher for longer-maturity portfolios and overall highest when BV is lower.

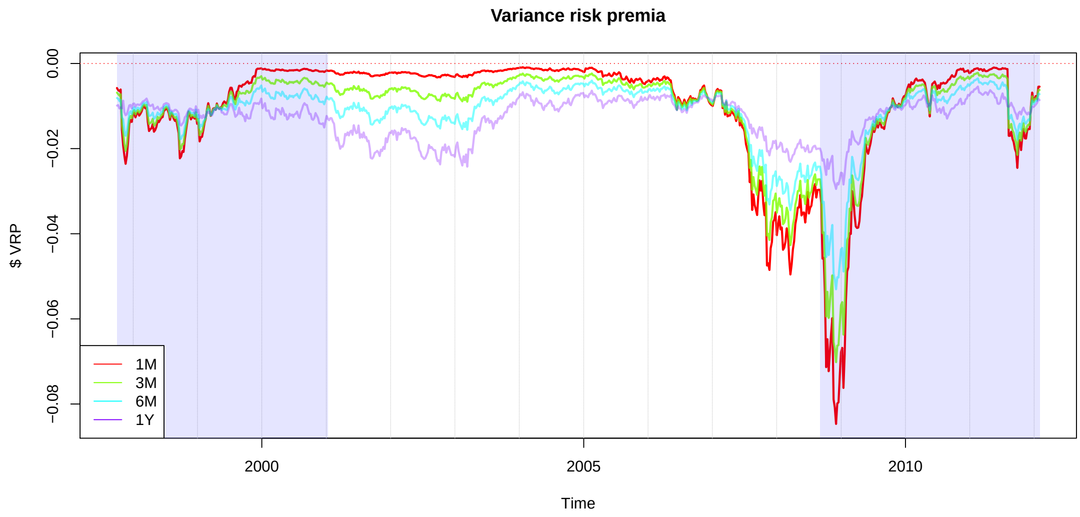


Figure 13: Dynamics of term structure of variance risk premia $\mathbb{E}_t^{\mathbb{P}}\left[QV_{t,t+h}\right] - \mathbb{E}_t^{\mathbb{Q}}\left[QV_{t,t+h}\right]$; blue shading denotes the out-of-sample period.

These results imply significant model misspecification, especially pertaining to periods of high market variation. The dynamics of returns on short-maturity options is captured by the model worse than this of longer-maturity options, which is particularly evident from strongly negative correlations the upper part of table 7.

	y	$q_{0.0} - q_{0.25}$	$q_{0.25} - q_{0.50}$	$q_{0.50} - q_{0.75}$	$q_{0.75} - q_{1.00}$	all
In sample	$\tau = 1M, u = -1$	0.073	-0.13	-0.1	-0.27	-0.038
	$\tau = 1M, u = 3$	0.097	-0.15	-0.08	-0.26	-0.046
	$\tau = 6M, u = -1$	0.16	0.12	0.16	-0.06	0.015
	$\tau = 6M, u = 3$	0.33	0.087	0.13	-0.015	0.027
Out of sample	$\tau = 1M, u = -1$	0.11	-0.048	-0.16	0.024	0.07
	$\tau = 1M, u = 3$	0.12	-0.077	-0.16	-0.014	0.05
	$\tau = 6M, u = -1$	0.12	0.11	0.029	0.11	0.093
	$\tau = 6M, u = 3$	0.089	0.28	0.015	-0.012	0.13

Table 7: Correlations of model-implied expected returns and realized returns, by subsample and stratified by spot BV quantiles.

6.2.3 Term structure of Variance Risk Premia

The rich dynamics of the term structure of variance risk premia is plotted in figure 13. Similarly to Gruber et al. [2015], Li and Zinna [2014] and Aït-Sahalia et al. [2012], the term structure is negative, downward-sloping in calm markets and upward-sloping in turbulent ones.

Compared to Aït-Sahalia et al. [2012] who fit variance swap rates, we obtain a richer dynamics of the shape of the TS: it switches slope significantly several times in 1998-1999, and then again in 2006 and 2009. Our VRP estimates are lower in magnitude in the earlier part of the sample but greater during the Financial Crisis.

Our risk premia estimates are of smaller magnitude than those of Li and Zinna [2014], especially in the aftermath of the Financial Crisis. Our estimates of the 1-year premium depend on the factors driving jumps only to a limited extent: throughout the sample it remains between 1 and 2.5%.

In the end of 2006 the TS switches sign and the variance risk premia increase in magnitude very significantly, especially at the short-end. Li and Zinna [2014] do not report such behavior. In Aït-Sahalia et al. [2012] the premia pick up in mid-2007, but it's the long end that moves first. Gruber et al. [2015] documents a move similar to the one we document, but of smaller magnitude.

The unconditional properties of the model-implied variance risk premia are summarised in table 8. The sample average VRP TS is flat and its standad deviation is decreasing with maturity. During the Financial Crisis the of the premium is up to 50% greater up to and including the 6-month maturity. The standard deviations of the VRP during and after the financial crisis are double those pre-crisis, again, up to the 6-month maturity.

6.2.4 Portfolio prices

If the model was correctly specified, return-based estimation would yield correct implied option portfolio prices. In this section we illustrate the difficulties that arise when pricing the chosen set of portfolios and we compare the results to an estimation which includes both portfolio returns and portfolio prices as inputs.

The reference return-based estimation produces model-implied portfolio prices that are markedly different from those observed in the market (panel

	Statistic	1M	2M	3M	6M	1Y
1997-2012	Mean	−0.0104	−0.0107	−0.011	−0.0118	−0.0128
	SD	0.0143	0.0127	0.0113	0.00802	0.00481
	Min.	−0.0847	−0.0771	−0.0702	−0.053	−0.0295
	$q_{0.05}$	−0.0399	−0.037	−0.0344	−0.0281	−0.0218
	$q_{0.5}$	−0.00389	−0.00542	−0.00688	−0.00969	−0.011
	$q_{0.95}$	−0.00121	−0.00214	−0.00293	−0.00482	−0.00742
	Max.	−0.000867	−0.00164	−0.00218	−0.00349	−0.00542
1997-2008	Mean	−0.00847	−0.0091	−0.00968	−0.0111	−0.0129
	SD	0.0108	0.00953	0.00845	0.00609	0.00455
	Min.	−0.0496	−0.046	−0.0427	−0.0345	−0.0242
	$q_{0.05}$	−0.0354	−0.0332	−0.0311	−0.0263	−0.0214
	$q_{0.5}$	−0.00307	−0.00514	−0.00671	−0.00978	−0.0114
	$q_{0.95}$	−0.00121	−0.00221	−0.00306	−0.00514	−0.00759
	Max.	−0.000867	−0.00164	−0.00235	−0.004	−0.00638
2008-2012	Mean	−0.0167	−0.016	−0.0155	−0.0141	−0.0121
	SD	0.021	0.0189	0.017	0.0121	0.00555
	Min.	−0.0847	−0.0771	−0.0702	−0.053	−0.0295
	$q_{0.05}$	−0.069	−0.0631	−0.0577	−0.0443	−0.0258
	$q_{0.5}$	−0.00816	−0.00839	−0.00861	−0.00923	−0.0104
	$q_{0.95}$	−0.00122	−0.00199	−0.0027	−0.00444	−0.00677
	Max.	−0.000939	−0.00167	−0.00218	−0.00349	−0.00542

Table 8: Variance risk premium: descriptive statistics; subsamples.

(a) of figure 14). If portfolio prices are included in the estimation dataset, the fit improves: visible mispricing occurs only during the Financial Crisis in 2009 (panel (b)). Further investigation of the properties of variance risk premia and return Sharpe ratios in the extended model reveals unsettling properties:

(i) The term structure of VRPs turns positive during the conundrum period and in September 2008; it is predominantly upward-sloping and changes its slope to *downward-sloping during the Financial Crisis*, contrary to most previous evidence.

(ii) Model-implied expected returns, variances and Sharpe ratios are further off from empirical values; Sharpe ratios at the longer end are significantly positive during the conundrum (thus are expected returns, even though the realized returns are negative); conditional return variances are underestimated.

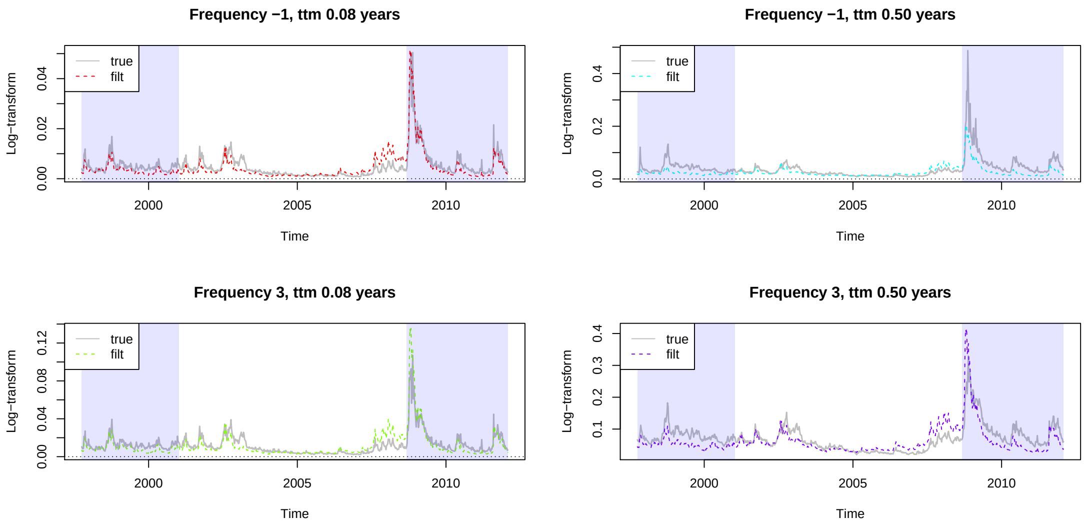
(iii) The variance jump risk premium is positive.

This preliminary evidence points to a very important consequence of attempting to fit both the dynamics of option portfolio returns and portfolio prices in a three-factor stochastic volatility model: the current specification is not sufficiently rich to capture the salient features of returns alongside prices.

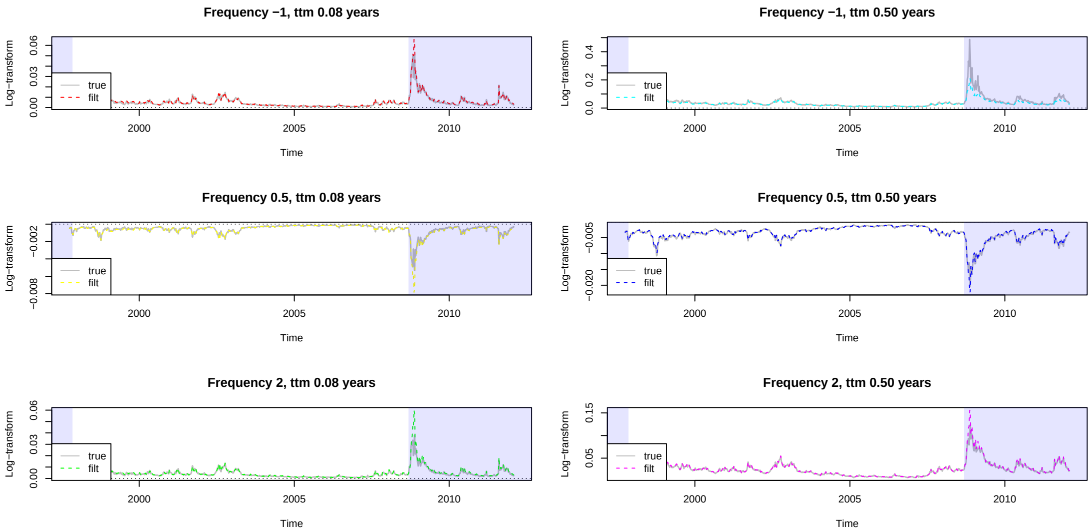
6.2.5 Portfolio price based estimation

For the sake of completeness we estimated the model using only log-transform values as inputs. In this case the $\mathbb{P}$-jump parameters are not identified. The estimation procedure was thus extended by treating filtered factor values as observable, calculating the conditional covariances of their increments with stock return increments and fitting $\mathbb{P}$-jump parameters via QMLE while keeping the other parameters fixed.

While a model thus estimated does a good job fitting portfolio values (see figure 15), a glance at the dynamics of the term structure of variance risk premia raises doubts about the model fit. The slope of the VRP almost doesn't change over the course of the sample, except for the Financial Crisis. Li and Zinna [2014] documented similar behaviour while using variance swap rates as inputs. Variance swap rates are – as logarithms of power portfolio values – linear combinations of variance factor values. The unconditional Sharpe ratios implied by the model (red lines) underestimate short-term Sharpe ratios and overestimate long-maturity Sharpe ratios (orange lines). The term structure of the conditional Sharpe ratios is also markedly different from what we obtained when estimating the model on power portfolio returns.

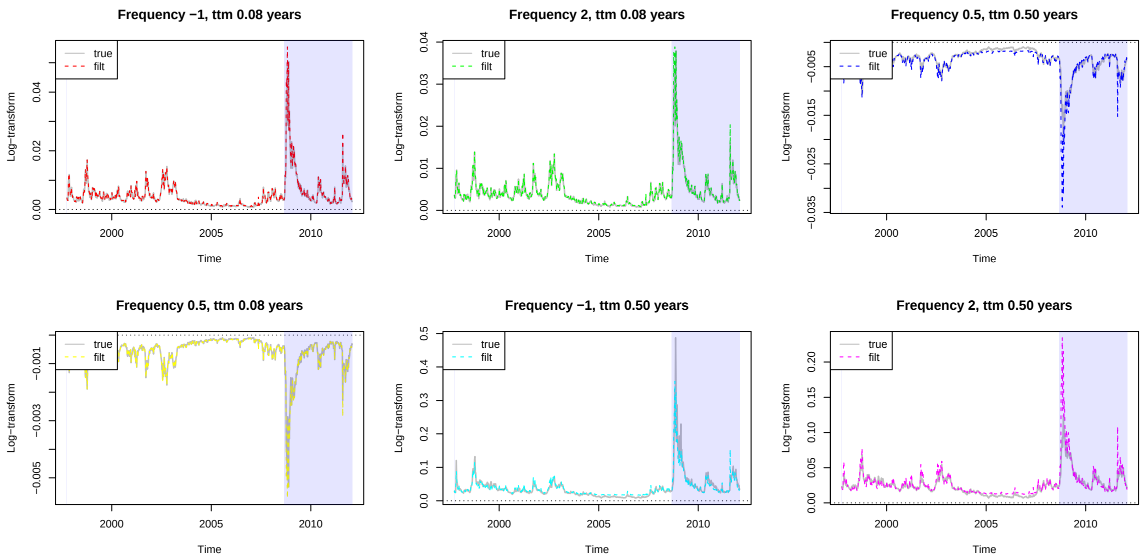


(a) Observed (grey) and fitted (coloured, dashed) prices of power portfolios from return-based estimation.

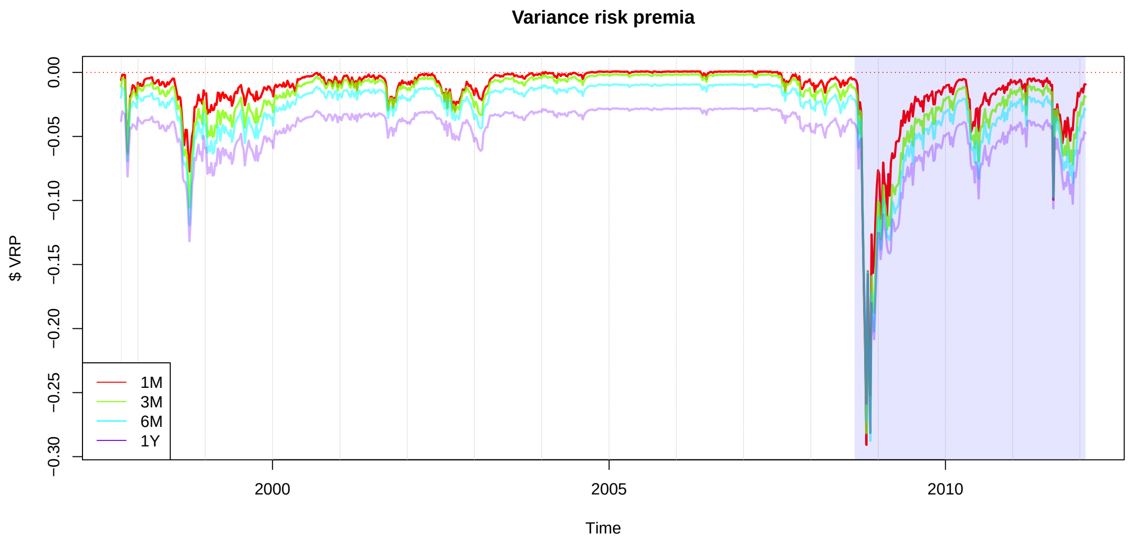


(b) Observed (grey) and fitted (coloured, dashed) prices of power portfolios from price-and-return based estimation.

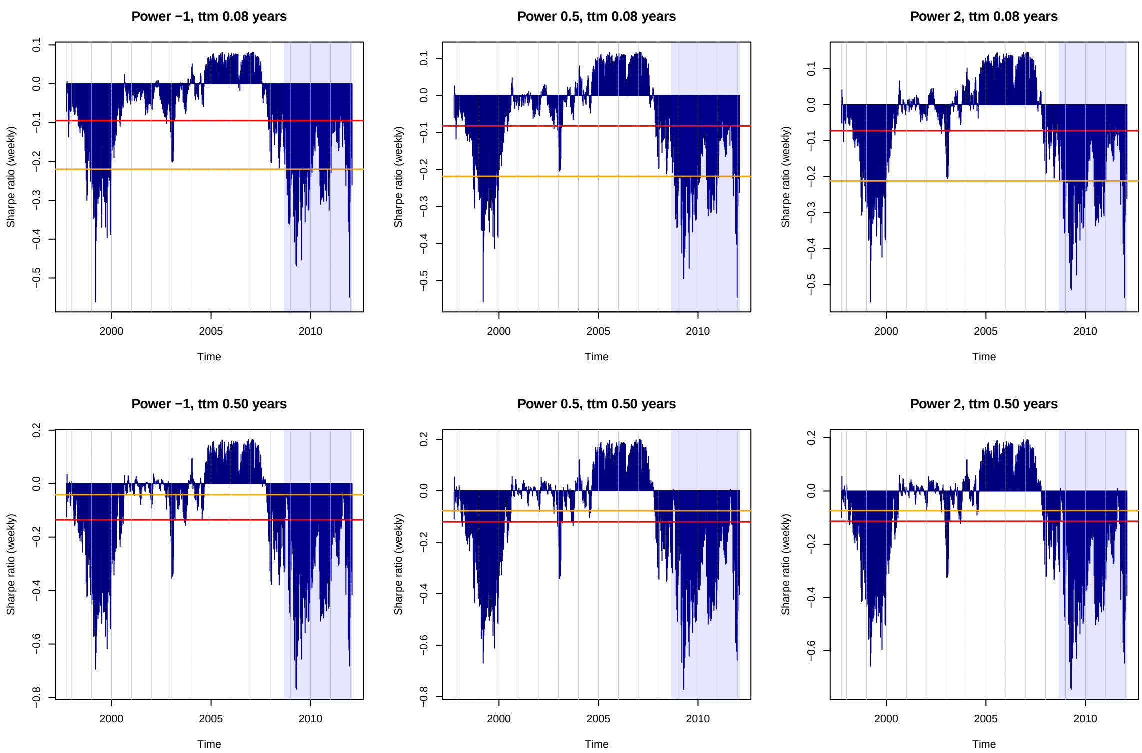
Figure 14: Observed and fitted option portfolio prices from two estimations.



(a) Observed (grey) and fitted (coloured, dashed) prices of power portfolios from return-based estimation.



(b) Term structure of variance risk premia.



(c) Predicted Sharpe ratios, orange line anchors the unconditional Sharpe ratio.

Figure 15: Results of model fit: price-based estimation.

7 Conclusions

Unspanned risks are the prime reason for the existence of option markets. Expected returns and variances of returns on portfolios yielding exposure to such risks are arguably of utmost interest for option investors. To this end, via proposition 1 we introduce a device for analytical treatment of these quantities in multifactor affine stochastic volatility models.

We show that using delta-hedged returns on power portfolios is beneficial for estimating unspanned-risk models. It directly exploits high-frequency information on the underlying asset. We have tentative evidence that under misspecification using the return instead of a price as an estimation target yields more reliable estimates of the stochastic properties of unspanned risks. We note that the implications for the dynamics of the term structure of variance risk premia are markedly different if returns are used for estimation rather than option prices. With the use of the Kálmán filter we propose a novel method of estimating jointly the complete $\mathbb{P}$ and $\mathbb{Q}$ dynamics of an option *risk* model.

In a Monte Carlo study we demonstrate the desirable properties of our approach, showing that the estimated conditional risk premia and risks are well captured both in the correctly specified and the misspecified cases, both in sample and out-of-sample. In contrast, traditional price-based estimation performed well only in the correctly specified case. Thus we believe that if the goal is dynamic risk premia estimation, then given the high probability of using a misspecified model, our approach should be favored. In addition, we also showed that we can create economically relevant gains in trading strategies by exploiting model implied conditional risk-premia estimates.

We contribute to understanding the risks associated with trading options. We document the time series of *returns* on a family of option portfolios which load on stochastic risk factors in varied ways. We note the high magnitude of the Sharpe ratios associated with these portfolios. The other important features of power portfolio returns are extremely high unconditional kurtosis and moderate right-skewness. We make the first attempt to estimate conditional Sharpe ratios of option portfolio returns and our results suggest that we need a richer model specification for attaining satisfying results. Nevertheless, our variance risk premium estimates exhibit the most important featues described in the literature to-date, such as: they are negative, persistent, and their unconditional term structure is flat, but its slope varies greatly over time.

A Appendix

A.1 Quadrature approximation

The first two conditional moments of $r_{t,h,\delta}$ defined in equation (16) can be calculated with the use of a highly accurate and computationally efficient approximation. In order to obtain it, first re-write the hedged return as:

(25)
$$r_{t,h,\delta} = \sum_{j=0}^{h/\delta} \frac{P_{t+j\delta,T-j\delta}}{P_{t,T-t}} \left(\frac{P_{t+(j+1)\delta,T-(j+1)\delta} - P_{t+j\delta,T-j\delta}}{P_{t+j\delta,T-j\delta}} - u\frac{F_{t+(j+1)\delta,T} - F_{t+j\delta,T}}{F_{t+j\delta,T}} \right)$$

(26)
$$=: \sum_{j=0}^{h/\delta} A_{t+j\delta,t+(j+1)\delta} \cdot \delta \,,$$

where $A_{t+j\delta,t+(j+1)\delta}$ is defined implicitly as the annualized hedged return increment over the period $[t + j\delta, t + (j + 1)\delta]$.

To calculate the conditional expectation of $r_{t,h,\delta}$ we need to calculate the sum:

(27)
$$\mathbb{E}_t^{\mathbb{P}}(r_{t,h,\delta}(u)) = \sum_{j=0}^{h/\delta} \mathbb{E}_t^{\mathbb{P}} \left[A_{t+j\delta,t+(j+1)\delta}\right] \cdot \delta =: \sum_{j=0}^{h/\delta} e_{t+(j+0.5)\delta} \cdot \delta \approx \int_0^{\delta} e(x)dx \,.$$

This can be seen as an approximation of the integral of the (smooth) function $e(x) := \mathbb{E}_t^{\mathbb{P}}(A_{t+x-\delta/2,t+x+\delta/2})$. However, with frequent hedging (small δ) the number of required evaluations becomes high. Therefore we evaluate the integral via Gaussian quadrature which requires much fewer approximation points:

$$\mathbb{E}_t^{\mathbb{P}}(r_{t,h,\delta}(u)) \approx \sum_{j=1}^{N} w_j e(x_j)\,, \tag{28}$$

where w_j are the Gauss-Legendre quadrature weights and x_j are the corresponding quadrature nodes. The required number of nodes depends on the mean-reversion speed of the variance factors. In practice, e is a smooth function of x, hence only a few quadrature points ($N \ll h/\delta$) suffice to calculate accurate approximations. In our calculations we found $N = 2$ quadrature points to be sufficient.

To calculate the conditional variance, we can use a similar approach, applying a two-dimensional quadrature method. In particular, we can write the squared hedged return as:

$$r_{t,h,\delta}^2 = \sum_{i=0}^{h/\delta}\sum_{j=0}^{h/\delta} A_{t+i\delta,t+(i+1)\delta}A_{t+j\delta,t+(j+1)\delta} \cdot \delta^2 = \tag{29}$$

$$= \sum_{i=0}^{h/\delta} A_{t+i\delta,t+(i+1)\delta}^2\delta^2 + \sum_{i\neq j;0\leq i,j\leq h/\delta} A_{t+i\delta,t+(i+1)\delta}A_{t+j\delta,t+(j+1)\delta} \cdot \delta^2 \,. \tag{30}$$

We separate the individual increment "variance" and "autocovariance" terms, since the former have different small-time behavior, due to the contemporaneity in the Brownian and jump components. We can thus write:

$$\mathbb{E}^{\mathbb{P}}(r_{t,h,\delta}^2) = \sum_{i=0}^{h/\delta} \mathbb{E}_t^{\mathbb{P}}[A_{t+i\delta,t+(i+1)\delta}^2\delta] \cdot \delta + \sum_{i\neq j;0\leq i,j\leq h/\delta} \mathbb{E}^{\mathbb{P}}[A_{t+i\delta,t+(i+1)\delta}A_{t+j\delta,t+(j+1)\delta}] \cdot \delta^2 \tag{31}$$

We use the approach from equation (28) to evaluate the first term. We notice that for $\delta \to 0$ the second term converges to the double integral of the function $v(x, y) = \mathbb{E}^{\mathbb{P}}[A_{t+x-\delta,t+x}A_{t+y,t+y+\delta}]$. Again, the smoothness of $v(x, y)$ allows for highly accurate sparse grid quadrature approximations, as in Petras [2003]. In practice, a second order accuracy (exact integration up to 3rd order polynomials) is sufficient, thus only 5 function evaluations are required.

Corollary 1 implies that the futures return can be expressed in the power portfolio framework, therefore the above approximation methods can be extended to calculating conditional means and covariances of the futures returns and the variance factors. This requires an approximation to variance factor increments: $(\exp(uV_{t+\delta}) - \exp(uV_t))u^{-1} \approx V_{t+\delta} - V_t$ for some small u (we choose $u = 10^{-3}$). The goal of using this approximation is to stay within the type of moments that can be calculated efficiently in the affine framework, as stated in proposition 1.

A.2 Empirical power portfolios

Our estimation method requires us to recover prices of power portfolios on a constant-maturity basis from panels of options of varying maturity. This section presents a method of recovering them from a panel of option prices on a given day. We also demonstrate how well our estimation method reproduces the theoretical prices of the power portfolios in our Monte-Carlo exercise in section 5.

Our approach uses all of the available option prices on any given day. First, we calculate Black-Scholes implied volatilities for all available option prices which are corrected for the existence of arbitrage opportunities[18]. Next, we fit a thin-plate regression spline to the squared values of the implied volatilities, where the two-dimensional surface $\mathcal{V}(\tau, \kappa)$ coordinates are time-to-maturity (τ) and log-moneyness scaled by the time-to-maturity ($\kappa := {}^{\log K}/_{\sqrt{\tau}}$). We fit the low rank Duchon [1977] isotropic spline using the procedure described in **?** and implemented in the package *mgcv* in R[19]. The thin-plate spline smoother can be considered a natural extension of univariate smoothing splines to higher dimensional data.

We choose the smoothing parameters $m = 1$ and $s = 0.5$ in the spline, which ensures that the extrapolating function is asymptotically constant. Such behaviour reduces the influence of the most out of the money options on the tails of the risk-neutral distribution. While this might seem limiting, the nature of the available data required us to restrict possible tail behavior. We note that in the simulated data the method performs admirably and allows to recover the known transform values from perturbed option prices with high accuracy.

Another advantage of fitting implied volatilites using regression splines is that the shrinkage target is meaningful. For example, if the penalty is of first order, then the shrinkage target is the Black-Scholes model, i.e. if the penalty parameter[20] $\lambda \to \infty$, then the fitted surface converges to a

[18] The correction procedure dwells on the fact that the differences in put prices at consecutive increasing strikes approximate the slope of the risk-neutral distribution function of the stock price, which should be non-decreasing.

[19] *mgcv* version 1.7-28, R version 3.0.2

[20] For the exact definition of the thin-plate optimization problem, we refer the reader to Hastie et al. [2009], Chapter 5.7 on Multidimensional Splines.

Black-Scholes model with constant volatility.

The prices of the power portfolios are then calculated for the requisite maturities by predicting the implied variance from the thin-plate spline, calculating corresponding Black-Scholes option prices and using the formulas given in Carr and Madan [2001] for replication.

A.3 Parameter estimates

This section contains the estimates of models under different specifications and methods.

	Parameter	Value	Parameter	Value
1	ρ_1^A	-0.2665	$\rho_2^{\mathbb{P}}$	-0.2603
2	ρ_3^A	-0.6598	$\lambda_{J,1}^{\mathbb{P}}$	0.4028
3	ϵ^{BV}	-6.8662	$\lambda_{J,2}^{\mathbb{P}}$	1.6618
4	ρ_1^C	0.8239	$\lambda_J^{\mathbb{P}}$	0.3187
5	ρ_3^C	0.6015	$\mu_J^{\mathbb{P}}$	-0.0168
6	ϵ_1^P	-8.1723	$\sigma_J^{\mathbb{P}}$	0.0430
7	ϵ_3^P	-9.8631	$\eta_1^{\mathbb{Q}}$	1.0193
8	ϵ_4^P	-8.6465	$\kappa_1^{\mathbb{Q}}$	13.0183
9	ϵ_6^P	-8.9272	$\eta_2^{\mathbb{Q}}$	1.1074
10	$\kappa_1^{\mathbb{P}}$	7.8208	$\kappa_2^{\mathbb{Q}}$	1.5100
11	$\sigma_1^{\mathbb{P}}$	3.8643	$\mu_J^{\mathbb{Q}}$	-0.0758
12	$\phi_1^{\mathbb{P}}$	0.1639	$\sigma_J^{\mathbb{Q}}$	0.1276
13	$\rho_1^{\mathbb{P}}$	-0.7622	ϵ_2^P	-8.8761
14	$\kappa_2^{\mathbb{P}}$	7.5264	ϵ_5^P	-8.8510
15	$\sigma_2^{\mathbb{P}}$	2.3425	ρ_2^A	-0.6387
16	$\phi_2^{\mathbb{P}}$	0.0830	ρ_2^C	0.8422

Table 9: Estimated parameters using the return-based calibration on the Monte Carlo sample of 500 weeks, using the incorrectly specified model (two factors).

	Parameter	Value	Parameter	Value
1	$\kappa_1^{\mathbb{P}}$	3.3365	$\mu_J^{\mathbb{Q}}$	-0.0125
2	$\sigma_1^{\mathbb{P}}$	1.2959	$\rho_J^{\mathbb{P}}$	-0.1171
3	$\phi_1^{\mathbb{P}}$	0.1764	$\sigma_J^{\mathbb{Q}}$	0.0737
4	$\rho_1^{\mathbb{P}}$	-0.6847	$\eta_1^{\mathbb{Q}}$	0.6382
5	$\kappa_2^{\mathbb{P}}$	4.1920	$\kappa_1^{\mathbb{Q}}$	3.0817
6	$\sigma_2^{\mathbb{P}}$	1.3188	$\eta_2^{\mathbb{Q}}$	1.0789
7	$\phi_2^{\mathbb{P}}$	0.0095	$\kappa_2^{\mathbb{Q}}$	3.0255
8	$\rho_2^{\mathbb{P}}$	-0.9999	$\nu_J^{\mathbb{Q}}$	0.4000
9	$\lambda_{J,1}^{\mathbb{P}}$	0.0001	ϵ_O	-3.4640
10	$\lambda_{J,2}^{\mathbb{P}}$	4.0000	$\mu_J^{\mathbb{P}}$	-0.0204
11	$\lambda_J^{\mathbb{P}}$	0.0000	$\sigma_J^{\mathbb{P}}$	0.0527
12	$\nu_J^{\mathbb{P}}$	0.0000		

Table 10: Estimated parameters using the price-based calibration on the Monte Carlo sample of 500 weeks, using the incorrectly specified model (two-factor).

	Parameter	Value	Parameter	Value
1	$\kappa_1^{\mathbb{P}}$	3.8631	$\lambda_J^{\mathbb{P}}$	0.0000
2	$\sigma_1^{\mathbb{P}}$	1.3806	$\nu_J^{\mathbb{P}}$	0.0000
3	$\rho_1^{\mathbb{P}}$	-0.2353	$\mu_J^{\mathbb{Q}}$	-0.0620
4	$\kappa_2^{\mathbb{P}}$	11.6587	$\rho_J^{\mathbb{P}}$	-0.0583
5	$\sigma_2^{\mathbb{P}}$	4.8219	$\sigma_J^{\mathbb{Q}}$	0.0972
6	$\phi_2^{\mathbb{P}}$	0.1410	$\eta_1^{\mathbb{Q}}$	0.7968
7	$\rho_2^{\mathbb{P}}$	-0.8614	$\kappa_1^{\mathbb{Q}}$	1.9083
8	$\kappa_3^{\mathbb{P}}$	2.9494	$\kappa_3^{\mathbb{Q}}$	0.7673
9	$\sigma_3^{\mathbb{P}}$	2.2564	ϵ_O	-3.6221
10	$\phi_3^{\mathbb{P}}$	0.0910	$\nu_J^{\mathbb{Q}}$	0.1422
11	$\rho_3^{\mathbb{P}}$	-0.3571	$\mu_J^{\mathbb{P}}$	-0.0217
12	$\lambda_{J,1}^{\mathbb{P}}$	3.2408	$\sigma_J^{\mathbb{P}}$	0.0457
13	$\lambda_{J,2}^{\mathbb{P}}$	0.5519		

Table 11: Estimated parameters using the price-based calibration on the Monte Carlo sample of 500 weeks, using the correctly specified model.

References

Y. Aït-Sahalia, M. Karaman, and L. Mancini. The term structure of variance swaps, risk premia and the expectation hypothesis. *SSRN Working Paper*, 2012.

T. Andersen, V. Todorov, and N. Fusari. The risk premia embedded in index options. *Journal of Financial Economics*, Forthcoming, 2015a.

T.G. Andersen, N. Fusari, and V. Todorov. Parametric inference and dynamic state recovery from option panels. *Econometrica*, Forthcoming, 2015b.

Gurdip Bakshi and Nikunj Kapadia. Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies*, 16(2):527–566, 2003.

Gurdip Bakshi, Nikunj Kapadia, and Dilip Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16 (1):101–143, 2003.

Chris Bardgett, Elise Gourier, and Markus Leippold. Inferring volatility dynamics and risk premia from the s&p 500 and vix markets. *Swiss Finance Institute Research Paper*, (13-40), 2013.

David S. Bates. Maximum likelihood estimation of latent affine processes. *Review of Financial Studies*, 19 (3):909–965, 2006.

T. Bollersev, George Tauchen, and H. Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies*, 22 (11):pp. 4463–4492, 2009. doi: doi: 10.1093/rfs/hhp008. URL `http://rfs.oxfordjournals.org/content/22/11/4463.full.pdf+html`.

Tim Bollerslev and Viktor Todorov. Tails, fears, and risk premia. *The Journal of Finance*, 66(6):2165–2211, 2011. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2011.01695.x. URL `http://dx.doi.org/10.1111/j.1540-6261.2011.01695.x`.

Tim Bollerslev and Jeffrey M Wooldridge. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11 (2):143–172, 1992.

Michael W Brandt and Ping He. Simulated likelihood estimation of affine term structure models from panel data. *Available at SSRN 885682*, 2006.

Mark Broadie, Mikhail Chernov, and Michael Johannes. Understanding index option returns. *Review of Financial Studies, 2009*, 2009.

P. Carr and L. Wu. Variance risk premia. *Review of Financial Studies*, 22(3):1311–1341, 2009.

P. Carr, H. Geman, D. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *Journal of Business*, 75:305–332, 2002.

Peter Carr and Roger Lee. Robust replication of volatility derivatives. In *PRMIA award for Best Paper in Derivatives, MFA 2008 Annual Meeting*, 2008.

Peter Carr and Dilip Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 2001.

Peter Carr and Liuren Wu. Leverage effect, volatility feedback, and self-exciting market disruptions. *Bloomberg Portfolio Research Paper No. 2009-03-FRONTIERS; AFA 2011 Denver Meetings Paper*, 2011.

John H. Cochrane, J. H. A mean-variance benchmark for intertemporal portfolio theory. *The Journal of Finance*, 69(1):1–49, 2014. ISSN 1540-6261. doi: 10.1111/jofi.12099. URL `http://dx.doi.org/10.1111/jofi.12099`.

Jin-Chuan Duan and Jean-Guy Simonato. Estimating and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting*, 13(2):111–135, 1999.

J. Duchon. *Construction theory of functions of several variables*, chapter Splines minimizing rotation-invariant semi-norms in Solobev spaces, pages pp. 85–100. Springer, 1977.

Gregory R Duffee and Richard H Stanton. Estimation of dynamic term structure models. *The Quarterly Journal of Finance*, 2(02), 2012.

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

Daniel Egloff, Markus Leippold, and Liuren Wu. The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis*, 45:1279–1310, 10 2010. ISSN 1756-6916. doi: 10.1017/S0022109010000463.

Eugene F. Fama and Kenneth French. Common risk factors in the returns on stocks and bonds. *Journals of Financial Econometrics*, 33:3–56, 1993.

Christian Gourieroux, Alain Monfort, and Alain Trognon. Pseudo maximum likelihood methods: Theory. *Econometrica: Journal of the Econometric Society*, pages 681–700, 1984.

Peter H. Gruber, Claudio Tebaldi, and Fabio Trojani. The price of the smile and variance risk premia. *Working Paper*, 2015.

Trevor Hastie, Robert Tibshirani, Jerome Friedman, T Hastie, J Friedman, and R Tibshirani. *The elements of statistical learning*, volume 2. Springer, 2009.

Junye Li and Gabriele Zinna. The variance risk premium: Components, term structures, and stock return predictability. Working Paper, September 2014.

Ian Martin. Simple variance swaps. Working Paper 16884, National Bureau of Economic Research, March 2012. URL `http://www.nber.org/papers/w16884`.

Jun Pan. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50, 2002.

Knut Petras. Smolyak cubature of given polynomial degree with few nodes for increasing dimension. *Numerische Mathematik*, 93(4):729–753, 2003.

M. Rockinger and M. Semenova. Estimation of jump-diffusion process vis empirical characteristic function. FAME Research Paper Series, 2005.

Joshua V. Rosenberg and Robert F. Engle. Empirical pricing kernels. *Journal of Financial Economics*, 64 (3):341–372, 2002.

Paul G. Schneider and Fabio Trojani. Divergence and the price of uncertainty. Working Paper, October 2015.

Zhaogang Song and Dacheng Xiu. A tale of two option markets: Pricing kernels and volatility risk. *Chicago Booth Research Paper*, (12-10), 2012.