

Improving the Filtering of Latent States Using Option Price Data

By

Samuel Léveillé

A Dissertation presented in partial fulfilment
of the requirements for the Degree of
Master of Science in Administration
(Financial engineering)

HEC Montréal

2020-07-07

© Samuel Léveillé, 2020

Remerciements

Tout d'abord, j'aimerais commencer par remercier ma superviseure Geneviève Gauthier pour son incroyable support tout au long de la réalisation de cet ouvrage. Par sa disponibilité, sa patience, ses explications simples des sujets les plus complexes et à certaines occasions, sa compréhension de mon incompréhension, elle sut rendre la création de ce mémoire l'élément marquant de mon parcours à la maîtrise. Côté quelqu'un véhiculant autant de passion pour son travail ne peut que transmettre celle-ci à son entourage et c'est pourquoi travailler avec elle fut l'une de mes meilleures expériences à HEC Montréal.

J'aimerais également remercier le corps professoral de HEC Montréal, notamment David Ardia, Michel Deneault et Christian Dorion pour leurs conseils et leurs réponses à mes questions tout au long de cet ouvrage. Également, je remercie Mohamed Jabir du LACED pour son aide à l'obtention des données financières et à l'utilisation du matériel informatique.

Je remercie mes collègues de HEC Montréal, particulièrement Rémi Galarneau-Vincent pour toutes les discussions que nous avons eu au cours de la réalisation de ce travail, ainsi que Gabrielle Trudeau pour les échanges d'idées et les précieux commentaires ayant grandement enrichi la qualité de cet ouvrage.

Un remerciement spécial à Marko Savor pour ses conseils qui m'ont orienté vers le programme d'ingénierie financière ainsi que pour son soutien durant ce parcours.

Je souhaite remercier mes parents, Jean-Marc Léveillé, Geneviève Desrochers et Louise Millette, pour leur incroyable soutien durant mes études. Finalement, je remercie mes amis pour leur support tout au long de mon parcours académique.

Sommaire

Les données d'options sont nécessaires afin de capturer les moments d'ordre supérieur des rendements boursiers. En utilisant des techniques modernes d'accélération par carte graphique, nous démontrons comment l'utilisation de rendements et de prix d'option dans l'étape de rééchantillonnage d'un filtre à particules permet un filtrage beaucoup plus précis des variables latentes d'un modèle à variance et intensité de saut stochastique. Cela nous apporte une performance largement accrue en évaluation d'options dû à une meilleure estimation des moments conditionnels d'ordre supérieur reflétés dans les prix. Finalement, nous explorons la prime de risque de marché en la décomposant en sa partie diffusive et de saut.

Abstract

Options data are mandatory to properly model the higher-order conditional moments of markets returns. Using modern GPU-acceleration, we demonstrate how including returns and options in the resampling step of a particle filter yield a much more precise filtering of the latent states variables for a model with time-varying variance and jump intensity. This results in an increased pricing performance due to a better estimation of the higher-order conditional moments reflected in options prices. Finally, we explore the composition of the equity risk premium by decomposing it in its diffusive and jump component.

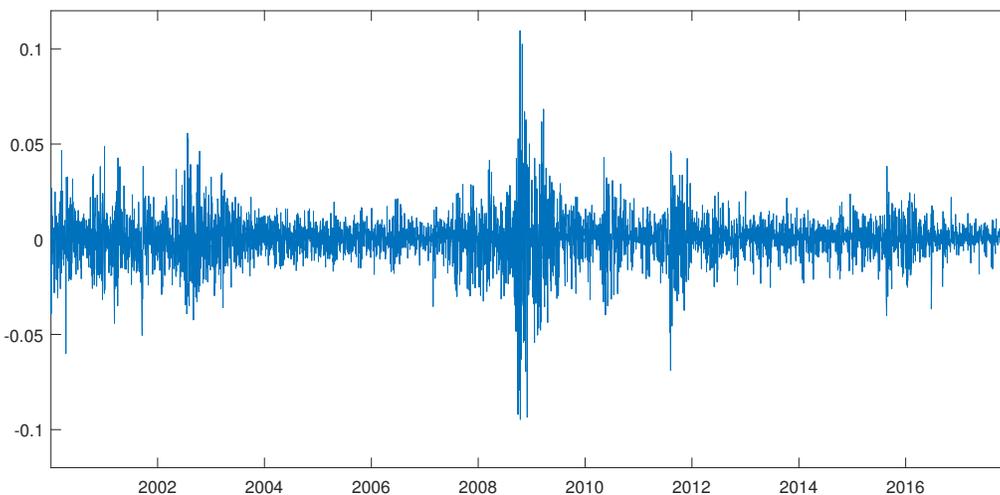
Contents

1	Introduction	5
2	Model	7
2.1	Under physical probability	7
2.2	Under risk-neutral probability	8
2.3	Option pricing formula	9
3	Model estimation	9
3.1	Data	10
3.2	Particle filter	10
3.3	Filtering with Returns	12
3.3.1	Parameter estimation using only Returns	12
3.3.2	Joint parameter estimation using Returns and Options	13
3.4	Filtering with Returns and Options	13
4	Simulation experiments	15
5	Empirical results	17
5.1	Estimated Parameters and Filtered latent state	17
5.2	Goodness of fit in the option market	18
5.3	Forecasting option prices	18
5.4	Risk-premium and returns conditional moments	19
6	Conclusion	21
7	Conclusion	22
A	APPENDIX	25
A.1	More on the model	25
A.2	Higher-order conditional moments	26
A.3	Option valuation	26
A.4	Numerical implementation	28
A.5	More on the Data	28

1 Introduction

Modelling security prices is of major importance in finance especially in option pricing. From the first model of Black and Scholes (1973) assuming a normal distribution for the underlying asset returns, models have improved in capturing the complex dynamics of financial time series with a literature separated in both continuous and discrete time framework. Looking at the daily returns of the S&P500 index in Figure 1, we observe that large fluctuations have been historically grouped together in periods of high volatility. This phenomenon of volatility clustering led to stochastic variance models such as Heston (1993) and Heston and Nandi (2000). However, this type of model focusing on time-varying volatility lack the flexibility to capture another characteristic of financial time series: returns are negatively skewed and punctuated with large fluctuations that traditional Gaussian innovations can't properly capture. This paved the way to models including jumps such as Bates (1996) and Ornthalalai (2014) composed of a diffusive and a jump innovation, with time-varying volatility and jump intensity. Today, most modern models include even more complex dynamics to better capture volatility persistence and extreme events. Notably, Andersen, Fusari, and Todorov (2015) used multiple volatility components, Amaya, Bégin, and Gauthier (2018) allowed for correlated jumps in both return and variance, and Bardgett, Gourier, and Leippold (2019) separated positive and negative jumps in returns.

Figure 1: Returns on the S&P500 index



Daily log-returns of the S&P500 index from January 1996 to December 2017.

The greatest challenge when using these complex models arises from the estimation of their parameters since relying only on returns observations fail to properly disentangle time-varying volatility, tail risk and risk premiums. This motivated the use of other sources

of information such as high-frequency data like Barndorff-Nielsen and Shephard (2004) who used realized variance and bipower variation in order to improve the estimation of diffusive coefficients. Another way to improve the estimation of models parameters is to add the cross-section of options prices to better capture extreme events, as well as risk premiums, since their prices provide information under risk-neutral probabilities.

A second complexity arises from the latent states variables present in most modern models that must be filtered. One of the popular approaches is the particle filter of Johannes, Polson, and Stroud (2009) that uses Monte Carlo simulation to extract latent states from observations. Although highly flexible and straight forward, the main drawback from this method is the increase in dimensionality of the filtering problem, resulting in a massive computational burden. This issue is particularly striking when options are also used in the estimation since computing their prices is already numerically costly, thus greatly limiting their uses. With the exceptions of Hurn, Lindsay, and McClelland (2015) who used parallel computing with two supercomputers to include options in the estimation of Heston (1993) model, Amaya, Bégin, and Gauthier (2018) using multiple sources of information¹ in their particle filter to estimate a jump-diffusion class of model, and Jacobs and Liu (2018) who proposed a novel solution based on model-implied spot volatilities rather than prices in order to reduce the dimensionality of the filtering problem, the most popular method consists in estimating the model’s parameters from a weighted likelihood of many information sources and using only returns to filter the latent states². While having the advantage of a reasonable computational time, this method doesn’t capture all of the information present in the options cross-section.

As models kept evolving over the years, another technology improved sidewise at a similar rate: Graphics Processing Units. Their use is widespread in artificial intelligence, bioinformatics, oil exploration, physics as well as many other scientific fields. In this paper, we tackled the computational burden of using options prices in a particle filter by the use of modern GPU-acceleration and explore the benefits of filtering latent states with returns and options data in order to estimate the parameters of time-varying volatility, tail risk and equity risk premiums.

Section 2 details the model we use, Section 3 the estimation procedure, Section 4 a simulation experiment and Section 5 the empirical result.

¹Amaya, Bégin, and Gauthier (2018) used returns, realized variance, bipower variation, options implied volatility and options realized variance to estimate a jump-diffusion model with correlated jumps in returns and variance

²This method was used notably by Christoffersen, Jacobs, and Ornathanalai (2012), Ornathanalai (2014) and Bégin, Dorion, and Gauthier (2020).

2 Model

To capture returns dynamics, we propose the discrete-time asset pricing model of Ornathanalai (2014) composed of a Brownian increment and a Lévy jump. Although not the most recent model, it allows us to study the filtering of time-varying variance, jump intensity and risk premium in parsimonious fashion.

2.1 Under physical probability

Returns R_{t+1} of the S&P500 index are modelled as

$$R_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) = \mu_{t+1}^{\mathbb{P}} - \zeta_{t+1} + z_{t+1} + y_{t+1}. \quad (1)$$

with the conditionally Gaussian and jump risk sources being z_{t+1} and y_{t+1} respectively. The term ζ_{t+1} is the convexity correction detailed in A.1 and $\mu_{t+1}^{\mathbb{P}}$ the drift under physical probability. Therefore, conditionally to the filtration \mathcal{F}_t^3 ,

$$\mathbb{E}^{\mathbb{P}}[S_{t+1}|\mathcal{F}_t] = S_t \exp(\mu_{t+1}^{\mathbb{P}}).$$

The conditionally Gaussian innovation follows

$$z_{t+1}|\mathcal{F}_t = \sqrt{h_{z,t+1}}\varepsilon_{t+1}$$

with ε_{t+1} being serially independent Gaussian increments. The conditional variance $h_{z,t+1}$ captures the volatility clustering under the affine GARCH dynamic of Heston and Nandi (2000)

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}} (z_t - c_z h_{z,t})^2.$$

The parameter w_z represents the long-term level of variance and c_z the asymmetric response to shocks in returns, described as the leverage effect by Black and Cox (1976). The unconditional level of variance is $\sigma_z^2 = \frac{w_z + a_z}{1 - b_z - a_z c_z^2}$ and the variance persistence is $b_z + a_z c_z^2$.

For the jump innovation, we choose the normal inverse Gaussian (NIG) distribution with a location parameter of 0, a tail heaviness of α , an asymmetry parameter of β and a time-varying scaling parameter $h_{y,t+1}$ referred to as the jump-intensity process⁴

$$y_{t+1}|\mathcal{F}_t \sim \text{NIG}(\alpha, \beta, 0, h_{y,t+1}).$$

The goal of jump innovations is to capture the skewness and heavy tail risk present in the index returns particularly for large negative jump and frequent small positive ones. The

³ $\mathcal{F}_t = \sigma\{z_{1:t}, y_{1:t}\}$, the σ -field generated by the diffusive and jump innovations from 1 to t .

⁴Bégin, Dorion, and Gauthier (2020) demonstrated that a NIG's scaling parameter can be seen as the scaled-jump intensity of a compound-Poisson process of NIG innovations.

NIG distribution was shown by Ornathanalai (2014) to have a better fit over the data than Merton's jumps and other Levy innovations⁵. The time-varying jump intensity dynamic is similar to the one of the variance process

$$h_{y,t+1} = w_y + b_y h_{y,t} + \frac{a_y}{h_{z,t}} (z_t - c_y h_{z,t})^2.$$

2.2 Under risk-neutral probability

Assuming the conditional Radon-Nikodym derivative that links the physical probabilities \mathbb{P} to the risk-neutral probabilities \mathbb{Q} is given by

$$\frac{\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_{t+1}}{\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t} = \frac{\exp(-\eta_{t+1}z_{t+1} - \Lambda_{t+1}y_{t+1})}{\exp(\xi_z(-\eta_{t+1})h_{z,t+1} + \xi_y(-\Lambda_{t+1})h_{y,t+1})}, \quad (2)$$

where $\xi_z(\phi)$ and $\xi_y(\phi)$ are linked to the cumulant generating function of the Gaussian and NIG innovations (see A.1). Prices discounted at the deterministic risk-free rate of returns $r_{f,t+1}$ under \mathbb{Q} are martingales

$$\mathbb{E}^{\mathbb{Q}} \left[e^{-r_{f,t+1}} \frac{S_{t+1}}{S_t} \middle| \mathcal{F}_t \right] = 1. \quad (3)$$

We apply the change of measure from the Radon-Nikodym derivative of Eq. (2)

$$\mathbb{E}^{\mathbb{Q}} \left[e^{R_{t+1}} \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{P}} \left[\frac{\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_{t+1}}{\frac{d\mathbb{Q}}{d\mathbb{P}}|\mathcal{F}_t} e^{R_{t+1}} \middle| \mathcal{F}_t \right] = \frac{\mathbb{E}_t^{\mathbb{P}}[\exp(-\eta_{t+1}z_{t+1} - \Lambda_{t+1}y_{t+1} + R_{t+1})]}{\exp(\xi_z(-\eta_{t+1})h_{z,t+1} + \xi_y(-\Lambda_{t+1})h_{y,t+1})}. \quad (4)$$

By replacing R_{t+1} in Eq.(4) with its definition from Eq.(1) and applying the martingale condition of Eq.(3) to the equation, we get

$$e^{r_{f,t+1}} = \frac{\mathbb{E}_t^{\mathbb{P}}[\exp(-\eta_{t+1}z_{t+1} - \Lambda_{t+1}y_{t+1} + \mu_{t+1}^{\mathbb{P}} - \zeta_{t+1} + z_{t+1} + y_{t+1})]}{\exp(\xi_z(-\eta_{t+1})h_{z,t+1} + \xi_y(-\Lambda_{t+1})h_{y,t+1})}.$$

Defining the convexity correction as $\zeta_{t+1} = \xi_z(1)h_{z,t+1} + \xi_y(1)h_{y,t+1}$ (see A.1) and rearranging the equation yield

$$e^{r_{f,t+1}} = \exp(\mu_{t+1}^{\mathbb{P}} - (\xi_z(1) + \xi_z(-\eta_{t+1}) - \xi_z(1 - \eta_{t+1}))h_{z,t+1} - (\xi_y(1) + \xi_y(-\Lambda_{t+1}) - \xi_y(1 - \Lambda_{t+1}))h_{y,t+1}).$$

⁵Ornathanalai (2014) demonstrated that the NIG distribution generate large levels of conditional non-normality while being parsimonious and has a coherent equity risk premium with observed market's returns

The pricing kernel allows for a decomposition of the instantaneous equity risk premium in terms of Gaussian and jump risk, assuming that η and Λ are deterministic and constant through time,

$$\mu_{t+1}^{\mathbb{P}} - r_{f,t+1} = \lambda_z h_{z,t+1} + \lambda_y h_{y,t+1}. \quad (5)$$

where

$$\begin{aligned} \lambda_z &= \xi_z(1) + \xi_z(-\eta) - \xi_z(1 - \eta) \\ \lambda_y &= \xi_y(1) + \xi_y(-\Lambda) - \xi_y(1 - \Lambda). \end{aligned}$$

The risk-neutral dynamic of returns thus follows

$$R_{t+1} = r_{f,t+1} - \zeta_{t+1}^* + z_{t+1}^* + y_{t+1}^*$$

with ζ_{t+1}^* being the convexity correction under \mathbb{Q} . The risk-neutral conditionally Gaussian innovation is $z_{t+1}^* | \mathcal{F}_t \sim N(0, h_{z,t+1}^*)$ and jump is $y_{t+1}^* | \mathcal{F}_t \sim \text{NIG}(\alpha, \beta^*, 0, h_{y,t+1}^*)$ with $\beta^* = \beta - \Lambda$. Under risk-neutral probabilities, the variance process and jump intensity leverage parameters are $c_z^* = c_z + \eta$ and $c_y^* = c_y + \eta$ respectively.

A complete proof of the risk-neutralization can be found in Ornathanalai (2014) and Bégin, Dorion, and Gauthier (2020).

2.3 Option pricing formula

Since our model selection fall into the exponential affine class, we can efficiently compute options prices by Inverse Fourier Transform for European options following the work of Heston and Nandi (2000). This aspect is crucial since our parameters estimation procedure rely on the computation of a large number of options prices. Hence, the price of a European call is

$$C_t(S_t, h_{z,t+1}, h_{y,t+1}) = S_t P_{1,t,T} - K e^{-r_{f,t,T}(T-t)} P_{2,t,T}.$$

with $r_{f,t,T}$ the deterministic risk-free rate for the corresponding maturity of the option. Details of the conditional probabilities $P_{1,t,T}$ and $P_{2,t,T}$ are available in A.3.

3 Model estimation

The latent model components z_t and y_t are estimated from the observed underlying asset returns and options prices via a particle filter. We compare two filtering methods, the first is based on index returns only and the second on index returns and the options cross-section.

We also compare three methods of parameter estimation. The first one is based on returns only, the second consist in filtering the latent states variables with returns and constructing a weighted likelihood function to include a sample of options in a two-stage method, and

the third one includes both returns and options prices in the filtering procedure to perform a one-stage estimation.

In all cases, the model’s parameters are determined via maximum likelihood estimation

$$\operatorname{argmax}_{\Theta} L(\Theta).$$

with Θ the vector of parameters⁶ being optimized

3.1 Data

To estimate the model, we use data from the S&P500 index between January 1,1996 and December 31, 2017, totalling 5,537 trading days. Excess index’s returns are calculated using data from CRSP and the 1-month T-bill rate from Fama and French library. The S&P500 options, dividend yields and zero-coupon rates are obtained from OptionMetrics. As it is now standard in the literature, a sample of the most liquid, out-of-the-money and at-the-money options contracts is selected, with moneyness⁷ between 0.75 and 1.15, and maturity between 10 and 90 days. Appendix A.5 details our selection procedure, resulting in a dataset of 47,186 options. For the zero-coupon rates associated with each contract, a linear interpolation is used in cases where the specific maturity isn’t available and a carry-over from the previous trading day is performed when exchanges are closed.

3.2 Particle filter

We describe a single step of a particle filter based on the sequential importance resampling (SIR) methodology. Conditioning on the jump paths $y_{1:t}^{(i)}$, where $i \in \{1,2,\dots,N\}$ and N is the number of particles simulated⁸, the conditional variance $h_{z,t+1}^{(i)}$ and jump intensity $h_{y,t+1}^{(i)}$ are predictable processes from information in time t . In our study, we propose two filtrations. The first is based on index returns where the sigma-field $\mathcal{G}_t^{(i)}$ is defined as

$$\mathcal{G}_t^{(i)} = \sigma\{R_{1:t}, y_{1:t}^{(i)}\}.$$

The second filtration includes the index returns as well as the options cross-section. The relevant sigma-field is

$$\mathcal{H}_t^{(i)} = \sigma\{R_{1:t}, \vec{O}_{1:t}^{mkt}, y_{1:t}^{(i)}\}$$

⁶The vector of parameters being $\Theta = \{w_z, a_z, b_z, \lambda_z, c_z, w_y, a_y, b_y, \lambda_y, c_y, \alpha, \beta\}$.

⁷We define moneyness as K/S_t .

⁸Throughout this paper, 25 000 particles were used.

with \vec{O}_t^{mkt} being the vector of M_t observed options prices

$$\vec{O}_t^{mkt} = \begin{bmatrix} O^{mkt}(S_t, K_1, \tau_1) \\ O^{mkt}(S_t, K_2, \tau_2) \\ \vdots \\ O^{mkt}(S_t, K_{M_t}, \tau_{M_t}) \end{bmatrix}.$$

Since $h_{z,t}^{(i)}$ and $h_{y,t}^{(i)}$ aggregate past information, $R_t, h_{z,t}^{(i)}, h_{y,t}^{(i)}, y_t^{(i)}$ is Markovian, which is a useful property for the filter. Therefore, the filtrations can be rewritten as

$$\mathcal{G}_t^{(i)} \equiv \sigma\{R_t, h_{z,t}^{(i)}, h_{y,t}^{(i)}, y_t^{(i)}\}.$$

and

$$\mathcal{H}_t^{(i)} \equiv \sigma\{R_t, \vec{O}_t^{mkt}, h_{z,t}^{(i)}, h_{y,t}^{(i)}, y_t^{(i)}\}$$

Step 1: Propagation

1. Since the deterministic risk-free rate of returns, the conditional variance and jump intensity are known at time t , we can directly compute the drift and convexity correction

$$\mu_t^{\mathbb{P}^{(i)}} = r_{f,t} + \lambda_z h_{z,t}^{(i)} + \lambda_y h_{y,t}^{(i)}$$

$$\zeta_t^{(i)} = \psi_{z_t}^{(i)}(1) + \psi_{y_t}^{(i)}(1).$$

2. The jump innovation is then simulated from the NIG distribution

$$y_t^{(i)} \sim \text{NIG}(\alpha, \beta, 0, h_{y,t}^{(i)}).$$

3. Once conditioned on the jump innovation, the conditionally Gaussian innovation is obtained

$$z_t^{(i)} = R_t - \mu_t^{\mathbb{P}^{(i)}} + \zeta_t^{(i)} - y_t^{(i)}.$$

4. Since the conditional variance and jump intensity are predictable processes in our framework, they are computed to be used in the next time step

$$h_{z,t+1}^{(i)} = w_z + b_z h_{z,t}^{(i)} + \frac{a_z}{h_{z,t}^{(i)}} \left(z_t^{(i)} - c_z h_{z,t}^{(i)} \right)^2$$

$$h_{y,t+1}^{(i)} = w_y + b_y h_{y,t}^{(i)} + \frac{a_y}{h_{z,t}^{(i)}} \left(z_t^{(i)} - c_y h_{z,t}^{(i)} \right)^2.$$

Step 2: Importance weights and filtering of latent state

1. Importance weights reflect the likelihood of a given particle and are based on the information available from the filtration, either $\bar{\omega}_t^{(i)}|\mathcal{G}_t^{(i)}$ or $\bar{\omega}_t^{(i)}|\mathcal{H}_t^{(i)}$. Since the main contribution of this paper is to compare the effects of both filtrations, we detail their computation in Equation (7) from Subsection 3.3 and Equation (11) from Subsection 3.4.

2. Normalized importance weights $\omega_t^{(i)}$ to be used in filtering and resampling are

$$\omega_t^{(i)} = \frac{\bar{\omega}_t^{(i)}}{\sum_{k=1}^N \bar{\omega}_t^{(k)}}. \quad (6)$$

3. Latent states variables are then filtered as the weighted average of observed particles

$$\begin{aligned} \tilde{z}_t &= \sum_{i=1}^N z_t^{(i)} \omega_t^{(i)} & \tilde{y}_t &= \sum_{i=1}^N y_t^{(i)} \omega_t^{(i)} \\ \tilde{h}_{z,t+1} &= \sum_{i=1}^N h_{z,t+1}^{(i)} \omega_t^{(i)} & \tilde{h}_{y,t+1} &= \sum_{i=1}^N h_{y,t+1}^{(i)} \omega_t^{(i)}. \end{aligned}$$

Step 3: Resampling

We resample the conditional variance and jump intensity based on the normalized weight of Eq. (6) following the 2-dimensional continuous resampling methodology from Malik and Pitt (2011):

$$h_{z,t+1}^{(i)} \leftarrow h_{z,t+1}^{(k_i)} \quad h_{y,t+1}^{(i)} \leftarrow h_{y,t+1}^{(k_i)}$$

where k_i are the particles being resampled. Since a multinomial resampling would yield a non-smooth likelihood function, this method allows us to achieve a likelihood function smooth enough for the use of conventional optimization routine.

3.3 Filtering with Returns

3.3.1 Parameter estimation using only Returns

This first estimation method, which will serve us as a benchmark, computes importance weights used in the resampling step of the particle filter based only on daily returns

$$\bar{\omega}_t^{(i)}|\mathcal{G}_t^{(i)} = f(R_t|\mathcal{G}_t^{(i)}). \quad (7)$$

Once the returns R_t are conditioned on the simulated jump path $y_{1:t}^{(i)}$, the Gaussian term can be isolated and its density is obtained

$$f(R_t|\mathcal{G}_t^{(i)}) = \frac{1}{\sqrt{2\pi h_{z,t}^{(i)}}} \exp\left(-\frac{1}{2} \frac{(R_t - \mu_t^{\mathbb{P}^{(i)}} + \zeta_t^{(i)} - y_t^{(i)})^2}{h_{z,t}^{(i)}}\right). \quad (8)$$

The likelihood function can then be computed as a direct by-product of the particle filter

$$L(\Theta) = L_{returns}(\Theta) = \sum_{t=1}^T \log \left(\frac{1}{N} \sum_{i=1}^N \bar{\omega}_t^{(i)} | \mathcal{G}_t^{(i)} \right). \quad (9)$$

3.3.2 Joint parameter estimation using Returns and Options

This second estimation procedure computes a weighted likelihood in two steps. First, the conditional variance and jump intensity processes are filtered using the importance weight computed from equation (7) and the log-likelihood from returns data $L_{returns}$ is obtained. The second steps computes the model's options prices based on filtered conditional variance and jump intensity

$$O_{t,m}^{mdl} = O \left(S_t, K_{t,m}, \tau_{t,m}, r_{f,t,m}, \tilde{h}_{z,t+1}, \tilde{h}_{y,t+1}, \Theta^{\mathbb{Q}} \right).$$

We then construct the likelihood function for options data using the Error Specification 4 (ES4) of Hurn, Lindsay, and McClelland (2015), a zero-mean multiplicative error with the constant σ_ε estimated as the error-term standard deviation

$$\varepsilon_{t,m} = \log(O_{t,m}^{mkt} / O_{t,m}^{mdl}) + \frac{1}{2} \sigma_\varepsilon^2 \sim N(0, \sigma_\varepsilon).$$

The density of each error being normally distributed

$$f \left(\varepsilon_{t,m} | \tilde{h}_{z,t+1}, \tilde{h}_{y,t+1} \right) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp \left(-\frac{1}{2} \frac{(\varepsilon_{t,m})^2}{\sigma_\varepsilon^2} \right),$$

we obtain the log-likelihood based on option observations with M_t being the amount of option included for each specific day,

$$L_{options}(\Theta) = \sum_{t=1}^T \sum_{m=0}^{M_t} \log(f \left(\varepsilon_{t,m} | \tilde{h}_{z,t+1}, \tilde{h}_{y,t+1} \right)). \quad (10)$$

A weighted likelihood combining the log-likelihood from returns (9) and from options (10) is then computed

$$L(\Theta) = \frac{T + M}{2} \left(\frac{L_{returns}(\Theta)}{T} + \frac{L_{options}(\Theta)}{M} \right).$$

with T the number for daily observation and M the number of options in the sample.

3.4 Filtering with Returns and Options

Our motive to include options data is to improve the filtering of latent states variables. From returns alone, it is impossible to distinguish between a large fluctuation in periods of high volatility and a jump during periods of low volatility. Jumps are also rare events, therefore

we can't achieve an accurate estimation of the jump intensity from returns only, especially if the intensity is time varying. Including the options cross-section to the filtering procedure solve these problems since prices directly reflect the conditional probability distribution of the underlying asset's returns. Also, Bégin and Gauthier (2020) found prices are highly biased when only historical returns are used to recover latent states and that this bias is corrected when option prices are added to the sample used to recover the states⁹.

Contrary to the two-stage estimation commonly used in the literature, this method values options prices based on the conditional variance and jump intensity of each particle instead of the filtered values

$$O_{t,m}^{mdl,(i)} = O \left(S_t, K_{t,m}, r_{f,t,m}, \tau_{t,m}, h_{z,t+1}^{(i)}, h_{y,t+1}^{(i)}, \Theta^{\mathbb{Q}} \right).$$

This is where most of the numerical challenge comes from since $N \times M$ options must be valued per likelihood calculation, close to a billion in our setting. This computational burden is only achievable by the use of advance numerical methods, see A.4 for a discussion from a numerical standpoint. Error terms are then computed for each particle following the same error specification as the two-stage method

$$\varepsilon_{t,m}^{(i)} = \log(O_{t,m}^{mkt} / O_{t,m}^{mdl,(i)}) + \frac{1}{2} \sigma_{\varepsilon}^2.$$

The error terms being normally distributed, the density is

$$f \left(\varepsilon_{t,m}^{(i)} | \mathcal{H}_t^{(i)} \right) = \frac{1}{\sqrt{2\pi\sigma_{\varepsilon}^2}} \exp \left(-\frac{1}{2} \frac{\left(\varepsilon_{t,m}^{(i)} \right)^2}{\sigma_{\varepsilon}^2} \right).$$

Assuming that errors from returns and options are independent, the importance weights conditional to the filtration $\mathcal{H}_t^{(i)}$ are computed as

$$\bar{\omega}_t^{(i)} | \mathcal{H}_t^{(i)} = f \left(R_t | \mathcal{H}_t^{(i)} \right) \prod_{m=0}^{M_t} f \left(\varepsilon_{t,m}^{(i)} | \mathcal{H}_t^{(i)} \right)^{1/M_t}, \quad (11)$$

with $f \left(R_t | \mathcal{H}_t^{(i)} \right)$ calculated the same way as $f \left(R_t | \mathcal{G}_t^{(i)} \right)$ in Eq. (8). To prevent options from overpowering returns observations, we weight options inversely proportional to the number of contracts observed that day, as proposed by Amaya, Bégin, and Gauthier (2018).

The log-likelihood is then computed from the importance weights of the particle filter

$$L(\Theta) = \sum_{t=1}^T \log \left(\frac{1}{N} \sum_{i=1}^N \bar{\omega}_t^{(i)} \right).$$

⁹In most models, options prices depend on unobservable factors such as stochastic volatility and jump intensity. Bégin and Gauthier (2020) demonstrated that the common practice of replacing latent states with their estimated values from returns induces a significant bias in options prices that is also propagated to models parameters.

Table 1: Summary of different filtration used in the simulation study

Filtration	Description
\mathcal{G}_t	A filtration containing only returns
\mathcal{H}_t^1	A filtration containing returns and one option per day. This option has a maturity of 30 days and a moneyness of 1.00.
\mathcal{H}_t^5	A filtration containing returns and five options per day. These options have a maturity of 30 days and a moneyness of 0.80, 0.90, 1.00, 1.10 and 1.20.
\mathcal{H}_t^{10}	A filtration containing returns and ten options per day. These options have a maturity of 30 and 90 days, and a moneyness of 0.80, 0.90, 1.00, 1.10 and 1.20.

A summary of the different filtrations used in our simulation experiment.

4 Simulation experiments

We propose a simulation study to illustrate the difference between the filtration \mathcal{G}_t and \mathcal{H}_t for the computation of importance weights in a controlled environment. Adding more options to the sample require a larger computational burden, therefore we also want to find the ideal number of options to maximize precision while keeping computation time reasonable.

We first simulate 2500 trading days under our model framework using parameters of the S&P500 from Bégin, Dorion, and Gauthier (2020), with a spot price starting at 1000, keeping zero-coupon rates and dividend yields at 0. We then simulate options prices for each day to be included in the filtering of the latent states. We add a multiplicative noise term to the simulated prices using the same error structure as in our estimation

$$O_{t,m}^{mkt} = O_{t,m}^{mdl} e^{-\frac{1}{2}\sigma_\varepsilon^2 + \varepsilon_{t,m}}$$

recalling that $\{\varepsilon_t\}_{t \in \mathbb{N}} \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon)$ and setting σ_ε to 0.05. To study the effect of different number of options included in the sample, we compare a filtration \mathcal{H}_t^1 containing one ATM option with 30 days to maturity, \mathcal{H}_t^5 containing 5 OTM/ATM options with 30 days to maturity and moneyness between 0.8 and 1.2, and \mathcal{H}_t^{10} containing 10 OTM/ATM options with 30 and 90 days to maturity and moneyness between 0.8 and 1.2. Table 1 summarize each filtration used.

We then run the particle filter using the same parameters as the data-generating-process. Table 2 present a comparative study using different performance measures. Panel A shows the average daily posterior standard deviation of the filtered process and Panel B the average error between the filtered process and the simulated one.

We find that information from the options cross-section vastly improves the filtering of latent states, mostly the conditional variance and jump intensity. The addition of one ATM option yields a reduction of the average posterior standard deviation of h_z and h_y , each by a factor of around threefold. When five options of the same maturity but different moneyness are

Table 2: Simulation results

Panel A: Average standard deviation of the posterior density				
	Filtering with Returns	Filtering with Returns and 1 Option	Filtering with Returns and 5 Options	Filtering with Returns and 10 Options
h_z	1.47E-05	3.63E-06	1.82E-06	1.75E-06
h_y	2.39E-05	6.48E-06	3.29E-06	3.17E-06
z	1.89E-03	1.74E-03	1.33E-03	1.29E-03
y	1.89E-03	1.74E-03	1.33E-03	1.29E-03

Panel B: Average error				
	Filtering with Returns	Filtering with Returns and 1 Option	Filtering with Returns and 5 Options	Filtering with Returns and 10 Options
h_z	-6.83E-06	-1.22E-06	1.00E-06	7.78E-07
h_y	-1.38E-05	-7.36E-06	-3.21E-06	-2.94E-06
z	-4.50E-04	-4.95E-04	-4.86E-04	-4.88E-04
y	-4.26E-04	-3.92E-04	-4.08E-04	-4.06E-04

Panel A shows the average standard deviation of posterior density of the filtered states. Panel B shows the average error, which is the difference between the filtered value and the simulated one.

used, we see a reduction of 7 and 6 times respectively. Interestingly, we observe that using 10 options, i.e. adding different maturities, improves but only slightly the filtering precision. Looking at the average error for h_z and h_y , we also see a significant improvement from the addition of options. For a more visual representation, the filtered values of h_z and h_y using \mathcal{G}_t and \mathcal{H}_t^{10} are presented in Figure 2.

For jumps innovation, we note a slight improvement in precision when using 5 or more options over returns alone but likely caused a more precise estimation of the jump intensity level itself. The improvement of the average error is insignificant. We graph the filtered jump innovations from \mathcal{G}_t and \mathcal{H}_t^{10} in Figure 3. From it, we observe that options help in filtering large jumps but still fail to capture small jumps innovations. A QQplot of the filtered Gaussian residuals is available in Figure 4, confirming an improvement in capturing large jumps when options are added to the filter.

From this experiment, we find that the addition of a panel of options greatly improves the filtering of variance and jump intensity but doesn't benefit the identification of jumps themselves. This directly validates our hypotheses that the addition of options to the filtration improve the estimation of h_z and h_y since their levels are directly reflected in options prices, but not y , as jumps innovations themselves don't affect prices in our model.

5 Empirical results

5.1 Estimated Parameters and Filtered latent state

Table 3 reports the model’s parameters values obtained from the three estimation methods. Beneath each parameter is the standard error computed using the outer product of gradients at the optimum parameters values (Newey and McFadden, 1994). From the Joint Estimation - Filtering with Returns, we observe similar parameters to Ornathanalai (2014) and Bégin, Dorion, and Gauthier (2020), which is to be expected since both papers used the same estimation methodology but on different datasets¹⁰.

Figure 5 and 6 present the filtered variance and jump intensity levels at the optimum parameters values for each estimation method. From the model filtered with returns and options, we first notice much tighter confidence intervals around the filtered values, indicating a gain in precision. The value of the parameter a_z of the Gaussian dynamic in Table 3 is the lowest off all the three estimation methods, indicating that the variance of the variance process is the lowest. This is confirmed by looking at the third plot of Figure 5. We also observe a larger leverage parameter c_y and a_y from the jump intensity dynamic, indicating a faster increase following negative returns. The third plot of Figure 6 shows filtering with returns and options yield the highest level of jump intensity during the financial crisis as compared to filtering only with returns. At the bottom of Table 3, we note that filtering with returns and options also yield the lowest level of unconditional volatility from Gaussian innovations. These findings indicate that extreme events are therefore more likely to be capture by jumps instead of large spikes in the variance level when options are used in the filtering of latent states.

The approach of variance targeting is popular in the financial literature (Mezrich and Engle, 1996) and consist in linking the value of a parameter to the variance of the market’s returns so that that the model’s unconditional variance is coherent with empirical observations¹¹. However, the presence of jumps in our modelling framework add a layer of complexity as both Gaussian and Jumps innovations contribute to the total variance. Therefore, we choose not to proceed with the method of variance targeting in our parameters search. Looking at the total unconditional volatility from both normal and jump innovations, at the bottom of Table 3, we see consistent values when compared to the sample’s volatility from market’s returns of 18.9%, validating that our methodological choice yield coherent results. The strong persistence of variance and jump intensity on the last row of Table 3 are also very similar to the results of Ornathanalai (2014) and Bégin, Dorion, and Gauthier (2020).

From the estimation using only returns, we notice in the first plot of Figure 6 that the jump

¹⁰Ornathanalai (2014) used data between 1996 and 2009, and Bégin, Dorion, and Gauthier (2020) between 1996 and 2015. Both only used Wednesdays options.

¹¹As an example, under the variance dynamic of Heston and Nandi (2000), the unconditionnal variance is calculated as $\sigma_z^2 = \frac{w_z + a_z}{1 - b_z - a_z c_z^2}$. Therefore, the method of variance targeting would fix the value of the long term variance’s level parameter $w_z = \hat{\sigma}_z^2(1 - b_z - a_z c_z^2) - a_z$, where $\hat{\sigma}_z^2$ is the returns sample’s variance.

intensity level is nearly flat. The parameters of jump intensity dynamic in Table 3 shows the highest value of w_y as well as very small a_y and c_y compared to both estimation methods using options. This leads us to believe that time-varying jump intensity is miss-identified in the absence of options. Much higher values for a_z and c_z in the variance dynamic are also observed, generating large spikes in the variance to capture extreme events that would normally be captured by jumps, as shown in the first plot of Figure 5.

5.2 Goodness of fit in the option market

To compare the pricing performance and goodness of fit from every estimation method, we compute the Relative Implied Volatility Root Mean Square error (RIVRMSE) for two subsets of options

$$RIVRMSE = \sqrt{\frac{1}{N} \sum_{j \in \mathcal{O}} \left(\frac{IV_j^{mld} - IV_j^{mkt}}{IV_j^{mkt}} \right)^2},$$

where $IV_{t,m}^{mld}$ is the implied volatility for a specific option under the Black-Scholes framework from the model and $IV_{t,m}^{mkt}$ the observed market's implied volatility. The first subset is the sample of options used in the estimation of the model's parameters. The results from each method are shown at the bottom of Table 3. The second subset consists of a broad selection of OTM/ATM options, with moneyness between 0.775 and 1.15, and maturities between 10 and 365 days. We detail the goodness of fit by moneyness and maturity in Table 4, and by time-period in Table 5.

Looking at the overall RIVRMSE for the 3 estimation methods, it is clear that the addition of options increase the performance of the model, most noticeably when options are included in the filter. It is worth mentioning that options with maturities above 90 days present slightly higher pricing errors since they were not included in the estimation of the model's parameters, but are still priced reasonably well. This in part validates the out-of-sample performance of our methodology.

5.3 Forecasting option prices

To assess the out-of-sample performance of our methodology, we use data from January 2018 to June 2019 that wasn't included in the estimation procedure to perform predictions of options prices for one and five days ahead. Using Monte Carlo simulation, we first forecast the underlying asset's spot price, conditional variance and jump intensity k days ahead¹² and used them as inputs in the option pricing formula. Risk-free rates and dividend yields are kept the same as in t for simplicity. We then compute the Black-Scholes implied volatility

¹²More precisely, $\{S_{t+k}^{(n)}, h_{z,t+k+1}^{(n)}, h_{y,t+k+1}^{(n)}\}$ based on information from \mathcal{G}_t or \mathcal{H}_t , where $k \in \{1,5\}$, the number of days ahead that are forecasted and $n \in \{1, \dots, N\}$, the scenario being simulated.

for each scenario and calculate the forecast expected value.

$$\widehat{IV}_{j,t+k}^{mld} = \frac{1}{N} \sum_{n=1}^N (IV_j^{(n),mld})$$

We then compare the realized implied volatilities to our expected forecasts using RIVRMSE.

$$RIVRMSE = \sqrt{\frac{1}{N} \sum_{j \in O} \left(\frac{\widehat{IV}_{j,t+k}^{mld} - IV_{j,t+k}^{mkt}}{IV_{j,t+k}^{mkt}} \right)^2},$$

Table 6 and 7 detail our out-of-sample pricing performance. From the overall valuation error of both one and five days ahead predictions, we again confirm the superior pricing performance of our estimation method. We also see the smallest increase in pricing errors relative to its in-sample counterpart, proving that filtering with returns and options results in the model’s parameters performing much better, even out-of-sample.

5.4 Risk-premium and returns conditional moments

We plot instantaneous risk premium from Eq.5 on index returns for each estimation method in Figure 7. When options are used in the filtering of latent states, we notice that the equity risk premium achieve the highest value of any estimation method at 36.3% during the financial crisis.

Table 8 summarize the distributions of equity risk premium over our observation period for each estimation method. With an annualized mean of 5.28%, we note the strongest volatility in the equity risk premium when options are used in the filtering of latent states with an annualized standard deviation of 4.99% as well as having a positive skewness of 2.61 and a kurtosis of 11.66. Those findings are consistent with the distribution obtained by Martin (2017) who used a nonparametric approach to estimate the lower bound of the risk premiums from options prices and argued that premiums are volatile, exhibit right skewness and heavy tails¹³.

The ability for the model to decompose the equity risk premium in a diffusive and jump component allows us to take a more detailed look at each source of risk. Looking at λ_z and λ_y , we see that including options in the sample either in the filter or in a weighted likelihood function achieve similar values. However, varying levels of risk from each estimation method yields different overall risk premiums for variance and jump intensity levels. Figure 8 plot the composition of equity risk premium from each source of risk for the three estimation methods. We observed that jump risk premium is the major component of the equity risk premium, particularly during periods of high volatility like the financial crisis.

¹³Between January 1996 and January 2012, Martin (2017) observed a lower bound for the equity risk premium, at a 1-Month horizon, with a mean of 5.00%, a volatility of 4.60%, a skewness of 4.03 and a kurtosis of 27.60.

We can also decompose the contribution of each source of risk to the overall variance. From Ornathanalai (2014), the model’s total conditional variance is affected by both the conditionally Gaussian and Jump risk

$$V^{\mathbb{P}}[R_{t+1}|\mathcal{F}_t] = \sigma_{t+1}^2 = h_{z,t+1} + \psi''_{y_{t+1}|\mathcal{F}_t}(0),$$

with $\psi''_{y_{t+1}|\mathcal{F}_t}(0)$ being the second derivative of the conditional cumulant generating function of jumps innovations evaluated in zero. When latent states are filtered with returns and options, we find that jump risk constitutes more than 61% of the equity risk premium although contributing to only 21% of the overall volatility in our framework, see Figure 9 for a visual representation. This support the theory that investors are much more averse to the jump component than the Gaussian one.

Another interesting aspect of our model is its ability to capture higher order conditional moments, see A.2 for details on their calculation. We plot the conditional skewness and kurtosis of daily index returns in Figure 10, and summarize their average levels in Table 9.

Although equity risk premium and volatility level are fairly consistent across estimation methods, we observe some major differences in the conditional skewness and kurtosis when options are included in the filtering of latent states. We first notice less volatile as well as lower levels of conditional skewness and kurtosis for the underlying index’s returns when options are also used to recover the latent states variables. By taking a look at the yearly evolution of pricing errors in Table 5 indicating the goodness of fit, we see the years that observed the most improvement in pricing performance (2005, 2006 and 2017) are also the ones observing the highest level of conditional skewness and kurtosis as shown in Figure 10. We therefore conclude that filtering with only returns fails to properly model the conditional asymmetry and tail risk present in the underlying asset’s returns and the addition of options to the filtering of latent states variables is necessary to properly capture those higher-order conditional moments.

6 Conclusion

We demonstrated that adding options data to the resampling step of a particle filter, while being numerically challenging, improve the estimation of the model's parameters significantly, especially those of the jumps dynamic. With the advance in computational capabilities allowed by modern GPUs, implemented on traditional programming languages, estimation method that where once deemed infeasible by numerical constraints should be reconsidered. With models incorporating richer dynamics to capture conditional asymmetry and tail risk of asset's returns, not including options prices in the filtering procedure will inevitably yield a miss identification of the model's parameters. An interesting element to explore would be in improving the filtering of multi-component variance dynamic since options offer information on the volatility term structure. We leave this subject open for future research.

7 Conclusion

Nous avons démontré que l'ajout des données d'options dans l'étape de rééchantillonnage d'un filtre particulière, bien que numériquement coûteux, améliore significativement l'estimation des paramètres du modèle. Avec les avancées technologiques en matière de calcul numérique permises par les cartes graphiques, implémentées dans des langages de programmation courants, nous croyons que le temps de calcul n'est plus le facteur limitant qu'il était auparavant. Avec des modèles incluant des dynamiques complexes pour capturer l'asymétrie et les événements extrêmes des données financières, ne pas inclure les prix d'options dans l'estimation des variables latentes mènera nécessairement à une mauvaise identification des paramètres animant leur dynamique. Un élément intéressant à explorer serait le filtrage de modèle à multiple composante de variance puisque les options offrent de l'information sur la structure à terme de la volatilité. Nous laissons ce sujet ouvert à de futures recherches.

References

- Amaya, Diego, Jean-François Bégin, and Geneviève Gauthier (2018). “Extracting Latent States from High-Frequency Option Prices”. In: *Available at SSRN 2975355*.
- Andersen, Torben G, Nicola Fusari, and Viktor Todorov (2015). “The risk premia embedded in index options”. In: *Journal of Financial Economics* 117.3, pp. 558–584.
- Bardgett, Chris, Elise Gourier, and Markus Leippold (2019). “Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets”. In: *Journal of Financial Economics* 131.3, pp. 593–618.
- Barndorff-Nielsen, Ole E and Neil Shephard (2004). “Power and bipower variation with stochastic volatility and jumps”. In: *Journal of financial econometrics* 2.1, pp. 1–37.
- Bates, David S (1996). “Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options”. In: *The Review of Financial Studies* 9.1, pp. 69–107.
- Bégin, Jean-François, Christian Dorion, and Geneviève Gauthier (2020). “Idiosyncratic jump risk matters: Evidence from equity returns and options”. In: *The Review of Financial Studies* 33.1, pp. 155–211.
- Bégin, Jean-François and Geneviève Gauthier (2020). “Price bias and common practice in option pricing”. In: *Canadian Journal of Statistics* 48.1, pp. 8–35.
- Black, Fischer and John C Cox (1976). “Valuing corporate securities: Some effects of bond indenture provisions”. In: *The Journal of Finance* 31.2, pp. 351–367.
- Black, Fischer and Myron Scholes (1973). “The pricing of options and corporate liabilities”. In: *Journal of political economy* 81.3, pp. 637–654.
- Christoffersen, Peter, Kris Jacobs, and Chayawat Ornthanalai (2012). “Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options”. In: *Journal of Financial Economics* 106.3, pp. 447–472.
- Heston, Steven L (1993). “A closed-form solution for options with stochastic volatility with applications to bond and currency options”. In: *The review of financial studies* 6.2, pp. 327–343.
- Heston, Steven L and Saikat Nandi (2000). “A closed-form GARCH option valuation model”. In: *The review of financial studies* 13.3, pp. 585–625.

- Hurn, A Stan, Kenneth A Lindsay, and Andrew J McClelland (2015). “Estimating the parameters of stochastic volatility models using option price data”. In: *Journal of Business & Economic Statistics* 33.4, pp. 579–594.
- Jacobs, Kris and Yuguo Liu (2018). “Estimation and Filtering With Big Option Data”. In: *Available at SSRN 3300564*.
- Johannes, Michael S, Nicholas G Polson, and Jonathan R Stroud (2009). “Optimal filtering of jump diffusions: Extracting latent states from asset prices”. In: *The Review of Financial Studies* 22.7, pp. 2759–2799.
- Malik, Sheheryar and Michael K Pitt (2011). “Particle filters for continuous likelihood evaluation and maximisation”. In: *Journal of Econometrics* 165.2, pp. 190–209.
- Martin, Ian (2017). “What is the Expected Return on the Market?” In: *The Quarterly Journal of Economics* 132.1, pp. 367–433.
- Mezrich, Joseph and RF Engle (1996). “GARCH for groups”. In: *Risk* 9.8, pp. 36–40.
- Newey, KW and D McFadden (1994). “Large sample estimation and hypothesis”. In: *Handbook of Econometrics, IV, Edited by RF Engle and DL McFadden*, pp. 2112–2245.
- Ornthanalai, Chayawat (2014). “Levy jump risk: Evidence from options and returns”. In: *Journal of Financial Economics* 112.1, pp. 69–90.

A APPENDIX

A.1 More on the model

The general modelling framework for index level under \mathbb{P} includes a drift $a^{\mathbb{P}}$, a Gaussian innovation z and a jump term y :

$$S_{t+1} = S_t \exp(a_{t+1}^{\mathbb{P}} + z_{t+1} + y_{t+1})$$

where S_t is the spot price of the underlying asset. The daily returns are calculated as

$$R_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right).$$

Under the physical probability, the conditional expected return at any time, assuming independent Gaussian and jump innovations is

$$\mathbb{E}^{\mathbb{P}}[S_{t+1}|\mathcal{F}_t] = S_t \exp(\mu_{t+1}^{\mathbb{P}})$$

but

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}[S_{t+1}|\mathcal{F}_t] &= S_t e^{a_{t+1}^{\mathbb{P}}} \mathbb{E}_t^{\mathbb{P}}[e^{z_{t+1}}] \mathbb{E}_t^{\mathbb{P}}[e^{y_{t+1}}] \\ &= S_t e^{a_{t+1}^{\mathbb{P}} + \psi_{z_{t+1}|\mathcal{F}_t}(1) + \psi_{y_{t+1}|\mathcal{F}_t}(1)} \end{aligned}$$

where $\psi_{X|\mathcal{F}_t}(\phi)$ is the conditional cumulant generating function evaluated in ϕ . Hence,

$$a_{t+1}^{\mathbb{P}} = \mu_{t+1}^{\mathbb{P}} - \psi_{z_{t+1}|\mathcal{F}_t}(1) - \psi_{y_{t+1}|\mathcal{F}_t}(1).$$

The conditional cumulant generating function of z_{t+1} is

$$\psi_{z_{t+1}|\mathcal{F}_t}(\phi) = \log(\mathbb{E}^{\mathbb{P}}[\exp(\phi z_{t+1})]) = \xi_z(\phi) h_{z,t+1}$$

with

$$\xi_z(\phi) = \frac{\phi^2}{2}$$

and the one for y_{t+1}

$$\psi_{y_{t+1}|\mathcal{F}_t}(\phi) = \log(\mathbb{E}^{\mathbb{P}}[\exp(\phi y_{t+1})]) = \xi_y(\phi) h_{y,t+1}$$

with

$$\xi_y(\phi) = \left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \phi)^2} \right).$$

Thus the convexity correction is

$$\zeta_{t+1} = \psi_{z_{t+1}|\mathcal{F}_t}(1) + \psi_{y_{t+1}|\mathcal{F}_t}(1).$$

A.2 Higher-order conditional moments

From Ornthalalai (2014), the conditional skewness and kurtosis of daily returns are

$$\text{Skew}^{\mathbb{P}}[R_{t+1}|\mathcal{F}_t] = \frac{\psi'''_{y_{t+1}|\mathcal{F}_t}(0)}{\sigma_{t+1}^{3/2}}$$

and

$$\text{Kurt}^{\mathbb{P}}[R_{t+1}|\mathcal{F}_t] = \frac{\psi''''_{y_{t+1}|\mathcal{F}_t}(0)}{\sigma_{t+1}^4}$$

with $\psi'''_{y_{t+1}|\mathcal{F}_t}(0)$ and $\psi''''_{y_{t+1}|\mathcal{F}_t}(0)$ being the third and fourth derivative of the conditional cumulant generating function evaluated in zero.

A.3 Option valuation

For $j \in \{1,2\}$, the conditional probabilities are

$$P_{j,t,T} = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[\frac{e^{-i\phi \log(\tilde{K})} \mathcal{M}_{R_{t,T}}^{\mathbb{Q}}(i\phi + 2 - j)}{i\phi} \right] d\phi \quad (12)$$

$$\tilde{K} = \frac{K e^{-rf_{t,T}(T-t)}}{S_t} \quad (13)$$

where Re is the real part of a complex number and $\mathcal{M}_{R_{t,T}}^{\mathbb{Q}}(\phi)$ is the conditional moment generating function of returns under \mathbb{Q} .

Following Ornthalalai (2014), the exponential affine modelling framework allows us to solve for the moment generating function of the asset price at any future time $t+k$, and in our case, the option maturity T .

$$\begin{aligned} \mathcal{M}_{R_{t,T}}^{\mathbb{Q}}(\phi) &= \mathbb{E}_t^{\mathbb{Q}}[S_{t+k}^\phi] = S_t^\phi \mathbb{E}_t^{\mathbb{Q}}[e^{\phi \sum_{i=1}^k R_{t+i}}] \\ &= S_t^\phi \mathbb{E}_t^{\mathbb{Q}}[\exp(A(\phi, t, t+k) + B(\phi, t, t+k)h_{z,t+1} + C(\phi, t, t+k)h_{y,t+1}^*)] \end{aligned}$$

Since at option maturity

$$\mathbb{E}_T^{\mathbb{Q}}[S_T^\phi] = S_T^\phi$$

The terminal values must be:

$$A(\phi, T, T) = B(\phi, T, T) = C(\phi, T, T) = 0$$

By iterated expectation

$$\begin{aligned}
\mathcal{M}_{R_t, T}^{\mathbb{Q}}(\phi) &= \mathbb{E}_t^{\mathbb{Q}}[\mathbb{E}_{t+1}^{\mathbb{Q}}[S_T^{\phi}]] = \mathbb{E}_t^{\mathbb{Q}}[\mathcal{M}_{R_{t+1}, T}^{\mathbb{Q}}(\phi)] \\
&= \mathbb{E}_t^{\mathbb{Q}}[S_{t+1}^{\phi} \exp(A(\phi, t+1, T) + B(\phi, t+1, T)h_{z, t+2} + C(\phi, t+1, T)h_{y, t+2}^*)] \\
&= S_t^{\phi} \mathbb{E}_t^{\mathbb{Q}}[\exp(\phi R_{t+1} + A(\phi, t+1, T) + B(\phi, t+1, T)h_{z, t+2} + C(\phi, t+1, T)h_{y, t+2}^*)]
\end{aligned}$$

Since the returns under \mathbb{Q} follow

$$R_{t+1} = r_{t+1} - \frac{1}{2}h_{z, t+1} + \sqrt{h_{z, t+1}}\varepsilon_{t+1}^* - \psi_{y_{t+1}|F_t}^*(1) + y_{t+1}^*, \quad (14)$$

we get

$$\begin{aligned}
\mathcal{M}_{R_t, T}^{\mathbb{Q}}(\phi) &= S_t^{\phi} \mathbb{E}_t^{\mathbb{Q}}[\exp(r_{t+1} - (\phi/2)h_{z, t+1} - \psi_{y_{t+1}|F_t}^*(1) + A(\phi, t+1, t+k))] \\
&\quad \times \mathbb{E}_t^{\mathbb{Q}}[\exp(\phi z_{t+1} + \phi y_{t+1}^* + B(\phi, t+1, t+k)h_{z, t+2} + C(\phi, t+1, t+k)h_{y, t+2}^*)].
\end{aligned}$$

By substituting the GARCH dynamic

$$\begin{aligned}
\mathcal{M}_{R_t, T}^{\mathbb{Q}}(\phi) &= S_t^{\phi} \mathbb{E}_t^{\mathbb{Q}}[\exp(\phi r_{t+1} - (\phi/2)h_{z, t+1} - \psi_{y_{t+1}|F_t}^*(1) + A(\phi, t+1, t+k))] \\
&\quad \times \mathbb{E}_t^{\mathbb{Q}}[\exp(\phi z_{t+1} + \phi y_{t+1}^* + B(\phi, t+1, t+k)h_{z, t+2} + C(\phi, t+1, t+k)h_{y, t+2}^*)] \\
&= S_t^{\phi} \mathbb{E}_t^{\mathbb{Q}}[\exp(r_{t+1} - (\phi/2)h_{z, t+1} - \phi \psi_{y_{t+1}|F_t}^*(1) + A(\phi, t+1, T) + B(\phi, t+1, T)(w_z + b_z h_{z, t+1}) + C(\phi, t+1, T)(w_y + b_y h_{y, t+1}^*))] \\
&\quad \times \mathbb{E}_t^{\mathbb{Q}}[\exp(\phi z_{t+1} + \phi y_{t+1}^* + B(\phi, t+1, t+k)(a_z/h_{z, t+1})(z_{t+1} - c_z^* h_{z, t+1})^2 + C(\phi, t+1, t+k)(a_y/h_{z, t+1})(z_{t+1} - c_y^* h_{z, t+1})^2)]
\end{aligned}$$

By using the properties:

$$\mathbb{E}_t^{\mathbb{Q}}[e^{\phi z_{t+1}}] = e^{(\phi^2/2)h_{z, t+1}} \quad \mathbb{E}_t^{\mathbb{Q}}[e^{\phi y_{t+1}^*}] = e^{\phi \psi_{y_{t+1}|F_t}^*(1)}$$

and

$$\mathbb{E}_t^{\mathbb{Q}}[e^{\alpha z_{t+1} + \beta z_{t+1}^2}] = e^{\alpha^2 h_{z, t+1}/2(1-2\beta h_{z, t+1}) - (1/2)\log(1-2\beta h_{z, t+1})},$$

we can isolate the following form of the moment generating function

$$\begin{aligned}
\mathcal{M}_{R_t, T}^{\mathbb{Q}}(\phi) &= S_t^{\phi} \exp(A(\phi, t, T) + B(\phi, t, T)h_{z, t+1} + C(\phi, t, T)h_{y, t+1}) \\
A(\phi, t, T) &= \phi r_{t+1} + A(\phi, t+1, T) + B(\phi, t+1, T)w_z + C(\phi, t+1, T)w_y \\
&\quad - \frac{1}{2}\log(1 - 2B(\phi, t+1, T)a_z - 2C(\phi, t+1, T)a_y) \\
B(\phi, t, T) &= -\frac{\phi}{2} + B(\phi, t+1, T)(b_z + a_z c_z^*) + C(\phi, t+1, T)a_y (c_y^*)^2 \\
&\quad + \frac{(\phi - 2B(\phi, t+1, T)a_z c_z^* - 2C(\phi, t+1, T)a_y c_y^*)^2}{2(1 - 2B(\phi, t+1, T)a_z - 2C(\phi, t+1, T)a_y^*)} \\
C(\phi, t, T) &= C(\phi, t+1, T)b_y - \phi \xi_y^*(1) + \xi_y^*(\phi)
\end{aligned}$$

The coefficients are solved recursively.

A.4 Numerical implementation

Re-Parametrization

The Nelder-Mead derivative-free algorithm was used for the parameters search since the rugged nature of the likelihood function made the numerical derivatives unreliable. This algorithm is an unconstrained optimization routine, therefore a re-parametrization was necessary to prevent the algorithm from taking steps outside the function domain, causing numerical instability. Also, this re-parametrization transform the discontinuous function into a continuous one, improving the search near discontinuities.

Parallelism

The estimation methods proposed in this study require a significant computational burden, particularly when options are included in the particle filter. Parallelism is the major factor in its realization. Since the computation of option prices at a specific time step can be divided into a large quantity of small independent tasks¹⁴, GPU-acceleration is the perfect tool for this application and massively outperform any traditional CPU-based implementation. Managing the memory of the GPU-device while also taking into consideration the bandwidth bottleneck from the PCIe was another numerical consideration. It should be noted that complex-double¹⁵ format limit the GPU processing performance significantly and should be avoided. In Table 10 we compared the computation time of the likelihood function for the Joint Estimation - Filtered with Return and Options when computed on a central processing unit (CPU) and a graphics processing unit (GPU).

Table 10: Likelihood computational time between a CPU and a GPU implementation

	Hardware	Specification	Computation time (Minutes)
CPU-based	Dual socket Intel Xeon E5 2667 V3	16 cores @ 3.20 GHz, 396 Go ECC DDR4	545.6
GPU-based	NVIDIA Quadro GP100	3584 CUDA Cores @ 1328 MHz, 16 Go ECC HBM2	11.0

We compare the computational speed of the GPU-accelerated implementation against one using only the CPU. Each likelihood function is evaluated once at the optimal parameter value. The GPU-device is installed in a system equipped with a Xeon W2195 processor (18 cores @ 2.3GHz). Both implementations are in MATLAB, without the explicit uses of lower-level programming languages.

The GPU-implementation achieve the same results 49 times faster than the CPU-based one. Considering the large amount of function valuation needed for the parameters estimation, this performance improvement made this method computationally feasible.

A.5 More on the Data

Here we detail the sampling procedure to obtain the options dataset used in the parameters estimation:

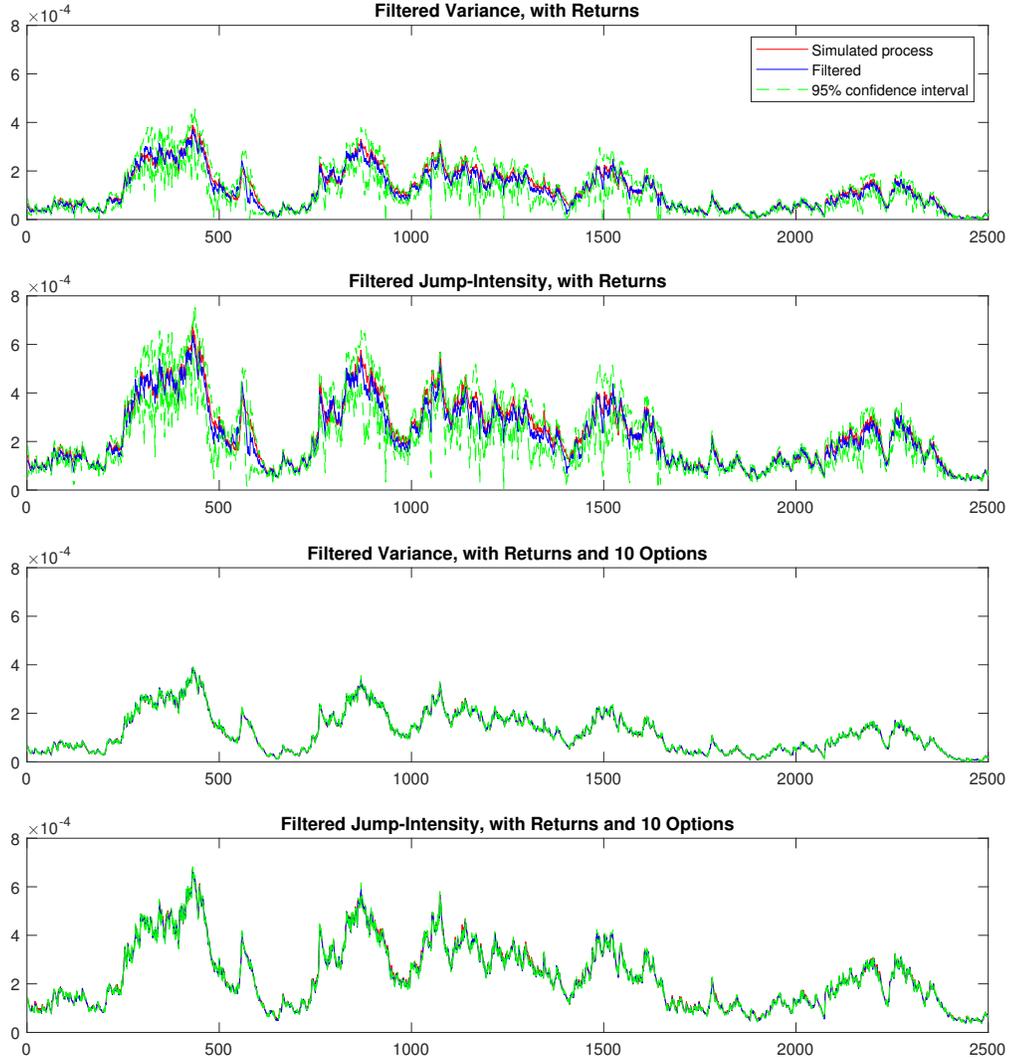
¹⁴We computed each integral by trapezoidal numerical integration using 1101 valuation points.

¹⁵Double-precision FP64 is needed since single-precision FP32 yield inaccurate results.

1. We first discard any contracts that suggests data errors such as $bid > ask$, Zero bid or not satisfying the non-arbitrage characteristic.
2. The second step is selecting contracts with positive Open-interest, a Volume above 50 and a mid prices above 0.375\$
3. We then made buckets by moneyness (K/S) ranging from 0.775 and 1.15 in steps of 0.075 and maturity of 10 to 45 and 45 to 90 days.
4. The most liquid option, OTM or ATM, from each bucket are then selected .

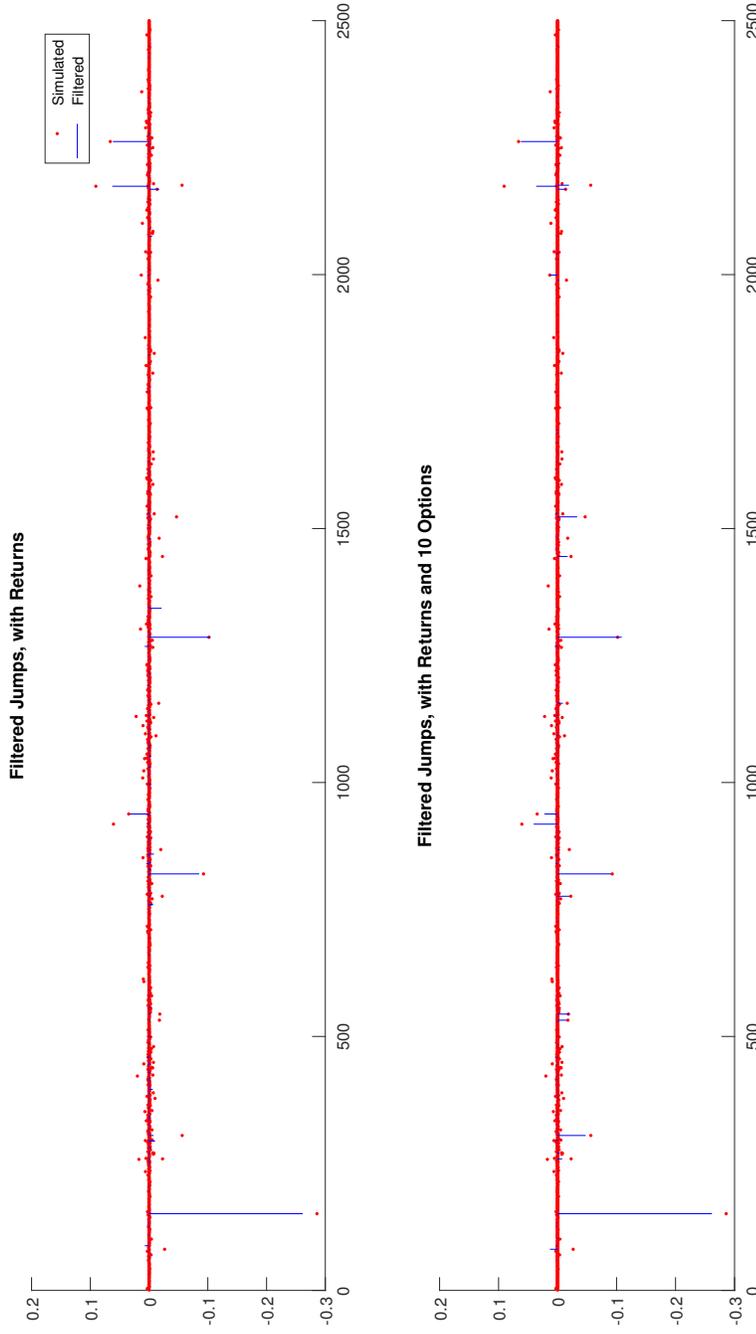
The final sample containt 47 186 contracts, with up to 10 per trading day. Table [11](#) give an overview of the final option sample.

Figure 2: Simulation result: filtered variance and jump-intensity



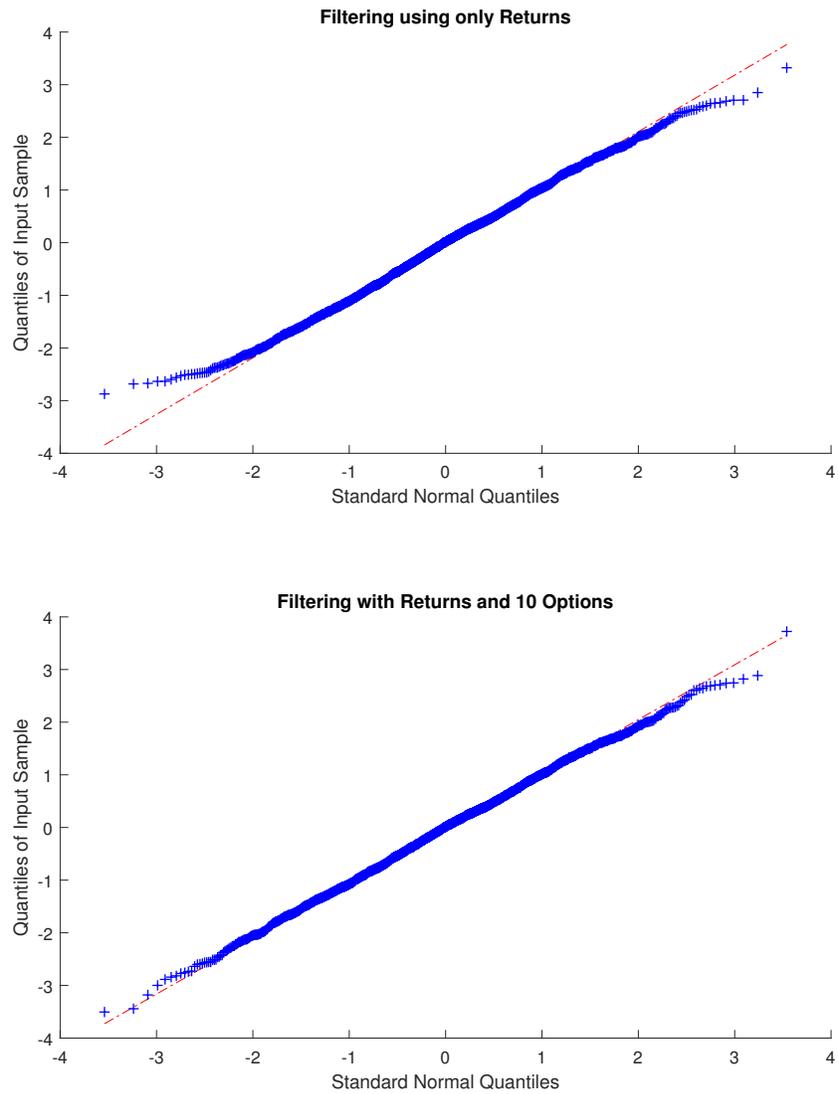
From the simulation on 2500 trading days, the conditional variance and jump intensity are filtered at the data-generating-process parameters values. The first two graphs reports the variance and jump intensity when filtered with \mathcal{G}_t and the bottom two when filtered with \mathcal{H}_t^{10} . The simulated values are plotted in red, the filtered values in blue and the 95% confidence intervals are shown in green.

Figure 3: Simulation result: filtered jump process



From the simulation on 2500 trading days, the jump process is filtered at the data-generating-process parameters values. The first graph reports jumps when filtered with \mathcal{G}_t and the bottom one with \mathcal{H}_t^{10} . The simulated values are in red and the filtered values in blue.

Figure 4: Simulation result: QQplot of Gaussian residues



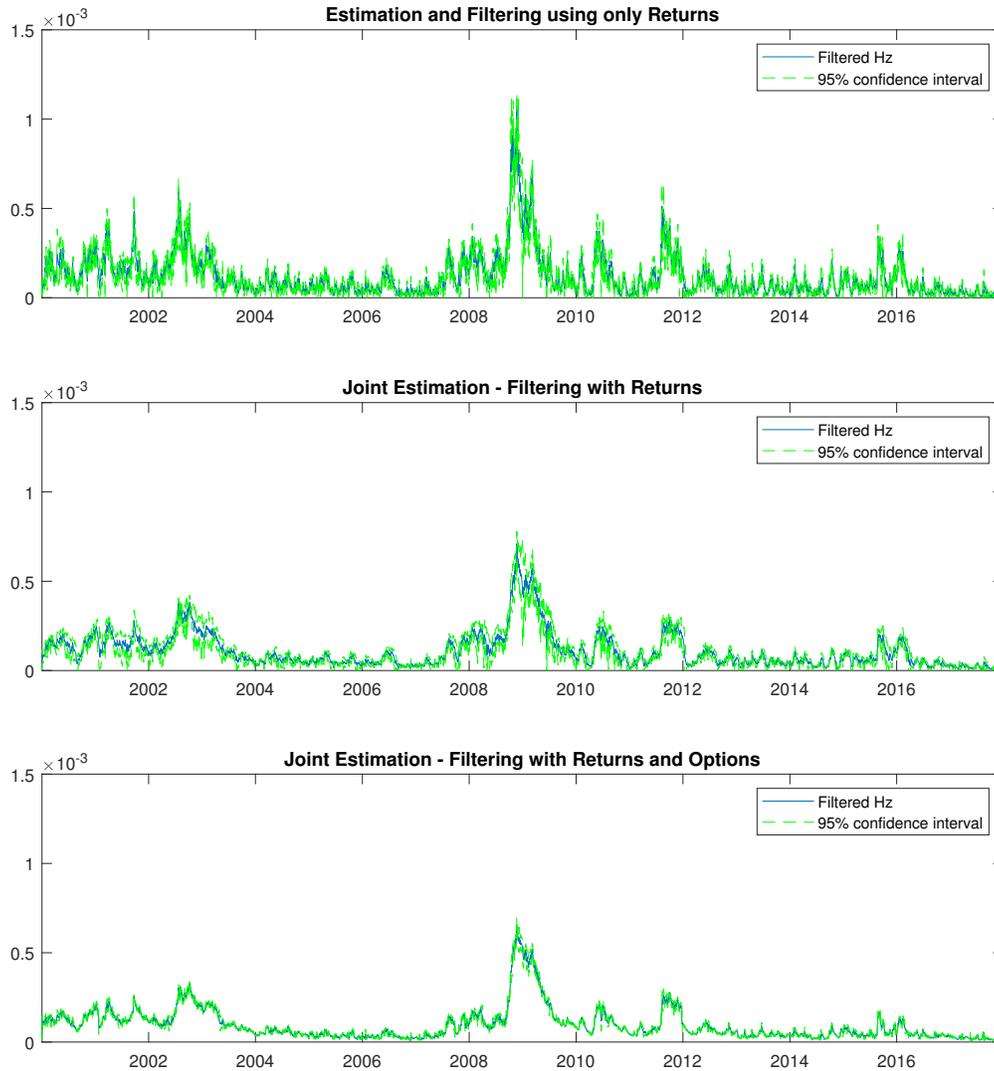
A QQplot of the Gaussian residuals from the simulation study. The first plot reports the Gaussian residues when filtered with \mathcal{G}_t and the bottom one with \mathcal{H}_t^{10} .

Table 3: Estimated parameter values on S&P500 from January 1996 to December 2017

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
<i>Gaussian</i>			
w_z	-1.95E-06 (3.25E-10)	-1.04E-06 (1.49E-09)	-6.34E-07 (1.40E-10)
a_z	5.16E-06 (1.55E-10)	2.26E-06 (3.26E-09)	1.62E-06 (2.16E-09)
b_z	0.814 (3.07E-04)	0.941 (9.26E-04)	0.945 (1.28E-03)
λ_z	0.770 (1.84E-03)	0.861 (2.13E-03)	0.838 (7.37E-04)
c_z	177.4 (2.84E-02)	149.3 (1.25E-01)	167.7 (2.45E-01)
<i>Jump</i>			
w_y	2.00E-05 (1.21E-08)	2.31E-07 (1.87E-09)	-1.16E-06 (5.72E-09)
a_y	2.91E-07 (9.46E-10)	2.59E-06 (4.52E-09)	4.09E-06 (9.70E-09)
b_y	0.901 (4.26E-05)	0.939 (1.78E-03)	0.943 (4.05E-04)
λ_y	0.525 (1.50E-04)	0.718 (4.55E-03)	0.678 (8.58E-04)
c_y	119.5 (2.31E-01)	140.3 (8.37E-01)	146.6 (5.64E-01)
α	14.41 (9.46E-03)	9.88 (2.44E-02)	13.17 (6.58E-02)
β	-8.60 (4.29E-03)	-5.92 (1.28E-02)	-7.32 (3.43E-04)
σ_ε	N/A	0.485	0.341
RIVRMSE	23.33	15.43	10.13
Log-Likelihood	14762.7	67306.1	12603.2
Uncond. Volatility (Gaussian)	18.37	18.76	16.19
Uncond. Volatility (Jumps)	8.44	9.03	8.37
Uncond. Volatility (Total)	20.22	20.82	18.22
Variance persistence	0.976	0.9913	0.9905

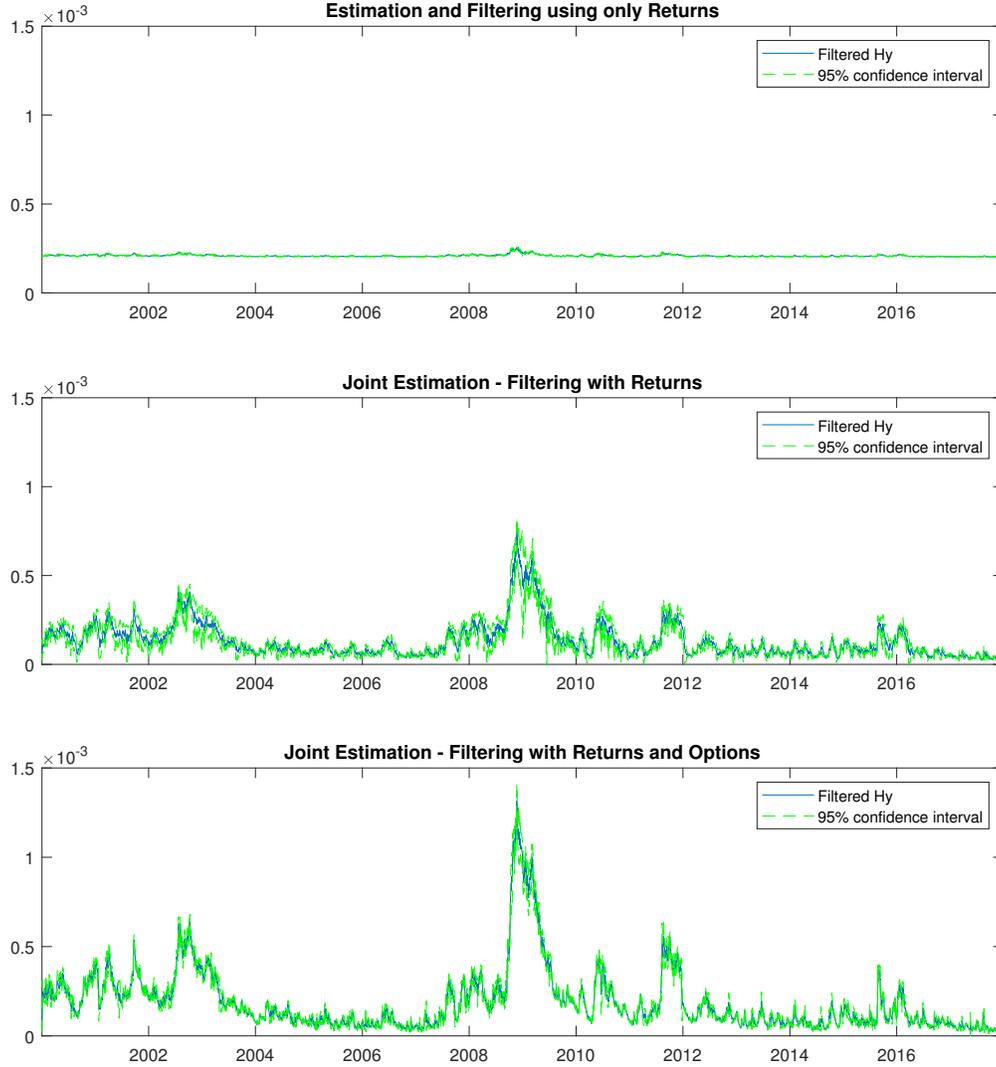
Parameters obtained from the maximum likelihood estimation on S&P500 from January 1996 to December 2017. Robust standard errors are computed from outer product of gradient at the optimal parameter values (Newey and McFadden, 1994). RIVRMSE and unconditional volatility are expressed in percentages. The total unconditional volatility is $\sqrt{252 \times (\sigma_z^2 + \sigma_y^2 \xi_y''(0))}$ where σ_z^2 is the unconditional variance and $\sigma_y^2 = (a_y + w_y + a_y c_y^2 \sigma_z^2) / (1 - b_y)$, detailed in Ornathanalai (2014).

Figure 5: Filtered conditional variance (h_z)



For the three methods of parameters estimation, we re-run the particle filter at the optimal parameters values to filter the conditional variance of the S&P500 daily returns. Confidence intervals are obtained from the posterior standard deviation of the filtered process. The first plot reports the filtered variance h_z when only returns are used in the parameters estimation, the second one when latent states are filtered with returns and parameters estimation is done with a two stages weighted likelihood, and the third one when returns and options are used to filter latent states. The filtered values are shown in blue and the 95% confidence intervals in green.

Figure 6: Filtered conditional jump intensity (h_y)



For the three methods of parameters estimation, we re-run the particle filter at the optimal parameters values to filter the conditional jump intensity of the S&P500 daily returns. Confidence intervals are obtained from the posterior standard deviation of the filtered process. The first plot reports the filtered jump intensity h_y when only returns are used in the parameters estimation, the second one when latent states are filtered with returns and parameters estimation is done with a two stages weighted likelihood, and the third one when returns and options are used to filter latent states. The filtered values are shown in blue and the 95% confidence intervals in green.

Table 4: Valuation error sorted by maturity and moneyness

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Overall	26.99	17.12	11.40
Sorted by day to maturity			
]10,45]	24.04	15.87	10.95
]45,90]	26.67	15.76	10.70
]90,180]	26.69	15.34	11.48
]180,270]	27.51	17.22	12.68
]270,365]	27.26	19.12	13.69
Sorted by moneyness			
]0.775,0.850]	14.02	9.92	10.19
]0.850,0.925]	17.81	10.74	10.43
]0.925,1.000]	26.23	13.71	10.27
]1.000,1.075]	36.38	22.85	13.82
]1.075,1.150]	31.92	24.99	15.26

Relative Implied-Volatility root mean square errors (RIVRMSE) are calculated from a panel of 538 832 options contracts on the S&P500 from January 2000 to December 2017. All values are expressed in percentages.

Table 5: Valuation error sorted by years

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Sorted by years			
2000	11.08	11.25	8.54
2001	10.45	10.24	7.97
2002	13.89	11.99	7.90
2003	12.03	8.20	7.39
2004	23.12	12.25	8.04
2005	42.88	25.47	11.72
2006	41.53	24.03	11.38
2007	28.31	17.98	12.12
2008	14.19	16.66	12.37
2009	20.66	11.36	11.33
2010	11.30	14.53	12.19
2011	13.15	13.66	12.30
2012	13.63	14.19	10.62
2013	25.57	13.54	7.72
2014	32.65	18.49	10.11
2015	28.37	23.59	12.39
2016	23.30	16.99	12.20
2017	45.40	19.24	15.05

Relative Implied-Volatility root mean square errors (RIVRMSE) are calculated from a panel of 538 832 options contracts on the S&P500 from January 2000 to December 2017. All values are expressed in percentages.

Table 6: Price prediction: one day ahead performance

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Overall	32.81	21.41	13.16
Sorted by day to maturity			
]10,45]	27.48	17.77	11.37
]45,90]	30.23	17.53	11.22
]90,180]	30.92	18.04	11.60
]180,270]	32.62	22.17	13.24
]270,365]	31.94	24.38	14.09
Sorted by moneyness			
]0.775,0.850]	9.48	6.42	5.74
]0.850,0.925]	15.68	8.03	5.77
]0.925,1.000]	28.89	14.93	7.94
]1.000,1.075]	48.00	32.54	20.19
]1.075,1.150]	46.11	38.21	22.91

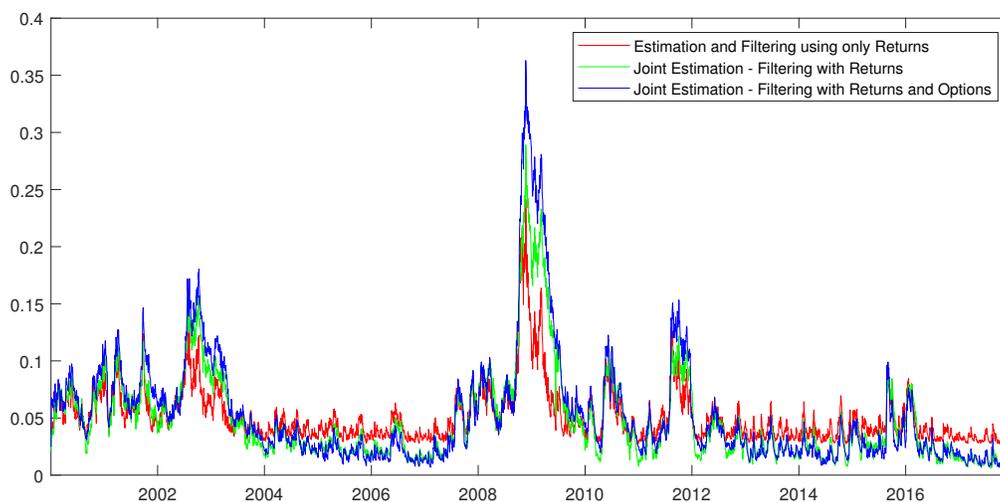
Parameters are estimated on the S&P500 from January 1996 to December 2017. Out-of-sample options are from January 2018 to June 2019 and are composed of a vast panel of OTM/ATM options. Prediction is done using Monte Carlo simulations. Performance is assessed by RIVRMSE using the predicted and realized options prices. All values are expressed in percentages.

Table 7: Price prediction: five day ahead performance

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Overall	34.07	23.98	15.18
Sorted by day to maturity			
]10,45]	29.75	18.63	12.13
]45,90]	32.07	18.49	11.94
]90,180]	32.27	18.92	12.24
]180,270]	33.35	23.01	13.88
]270,365]	32.50	25.08	14.68
Sorted by moneyness			
]0.775,0.850]	9.82	6.46	5.47
]0.850,0.925]	16.27	8.07	5.44
]0.925,1.000]	30.08	15.45	8.24
]1.000,1.075]	50.77	34.77	22.33
]1.075,1.150]	46.68	39.33	23.94

Parameters are estimated on the S&P500 from January 1996 to December 2017. Out-of-sample options are from January 2018 to June 2019 and are composed of a vast panel of OTM/ATM options. Prediction is done using Monte Carlo simulations. Performance is assessed by RIVRMSE using the predicted and realized options prices. All values are expressed in percentages.

Figure 7: Equity risk premium - Three estimation methods



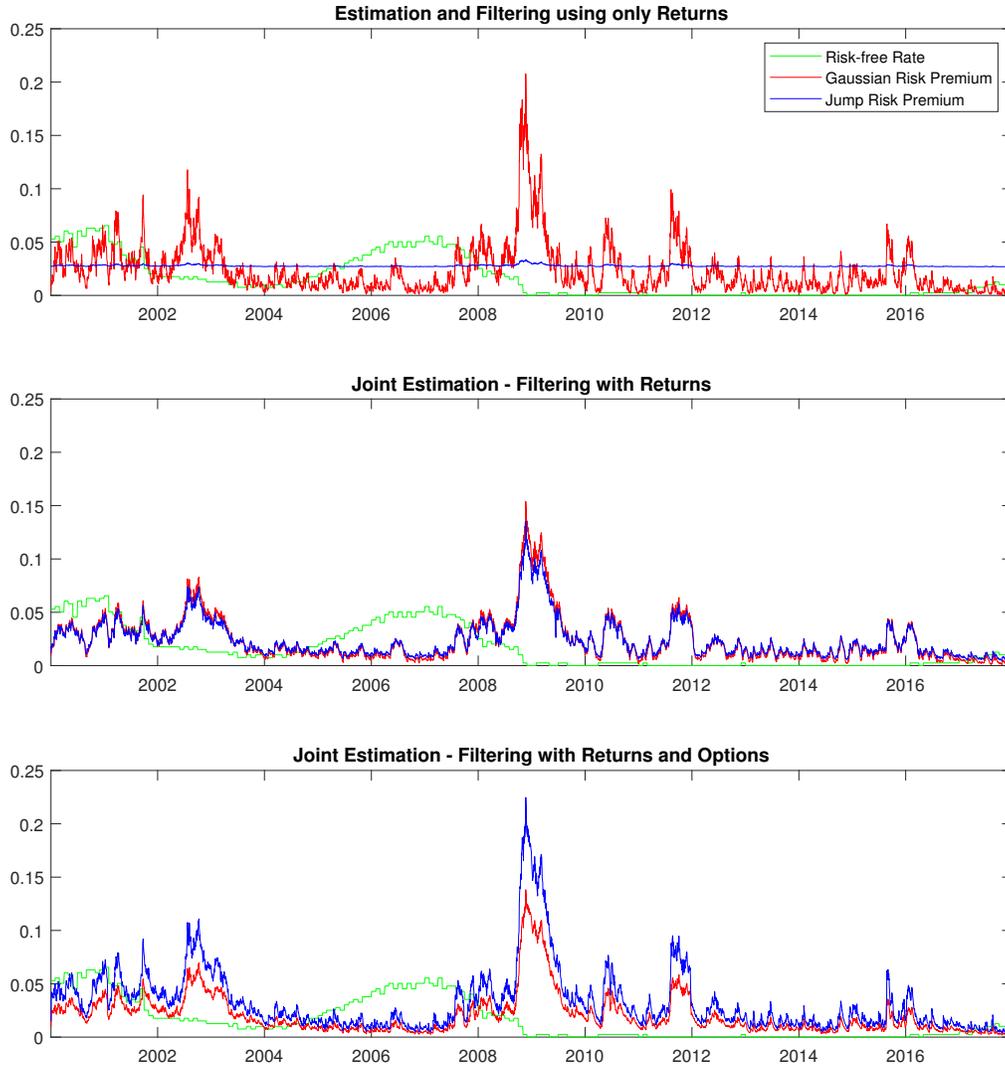
From the filtered conditional variance and jump intensity, we compute the equity risk premium (ERP) on the daily S&P500 index's returns between January 2000 to December 2017, for each of the three methods of estimation.

Table 8: Equity risk premium

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Average	5.10%	4.82%	5.28%
From diffusion	2.33%	2.40%	2.03%
From jumps	2.77%	2.41%	3.25%
standard deviation	2.40%	3.95%	4.99%
skewness	2.71	2.16	2.61
kurtosis	13.85	9.10	11.66
Minimum	2.71%	0.48%	0.57%
Maximum	24.15%	28.91%	36.30%

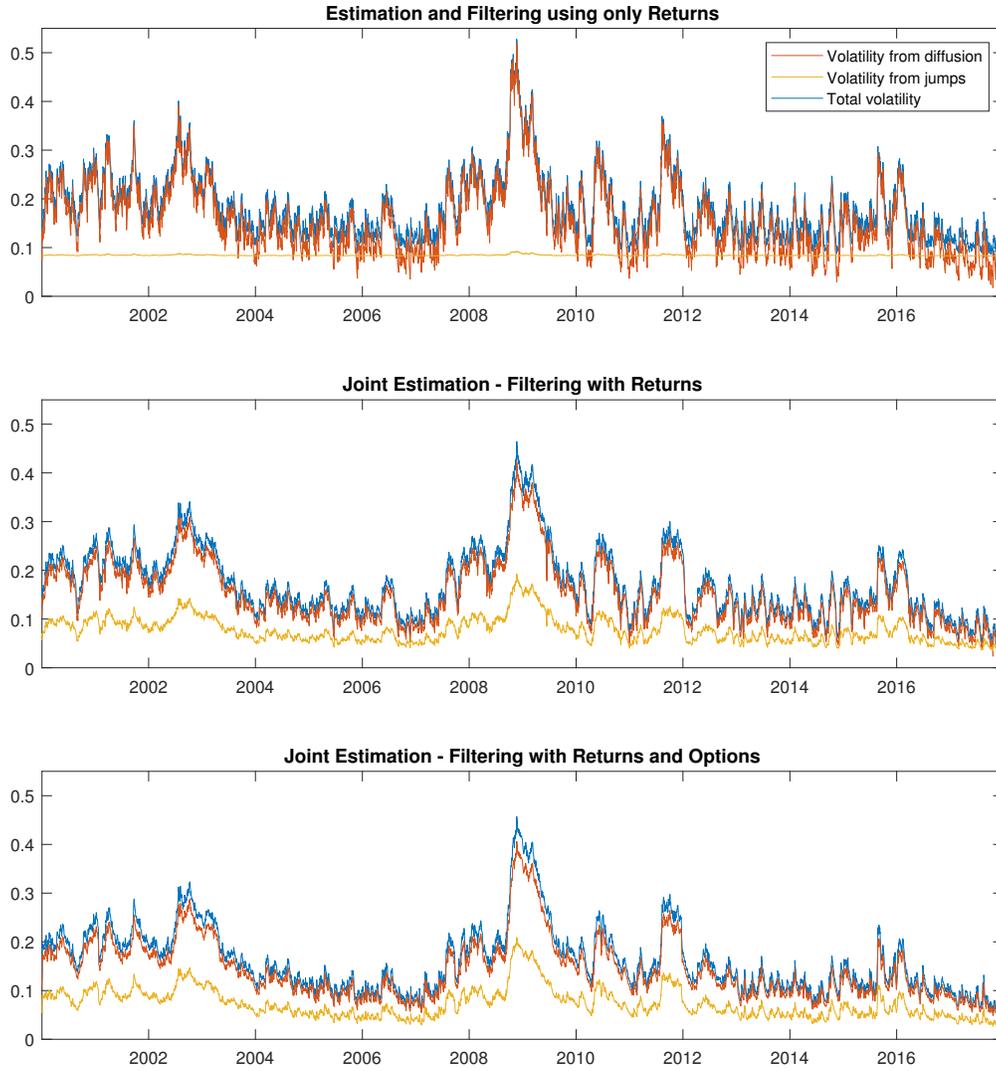
Summary of the filtered annualized equity risk premium for the daily S&P500 index's returns between January 2000 to December 2017.

Figure 8: Equity risk premium



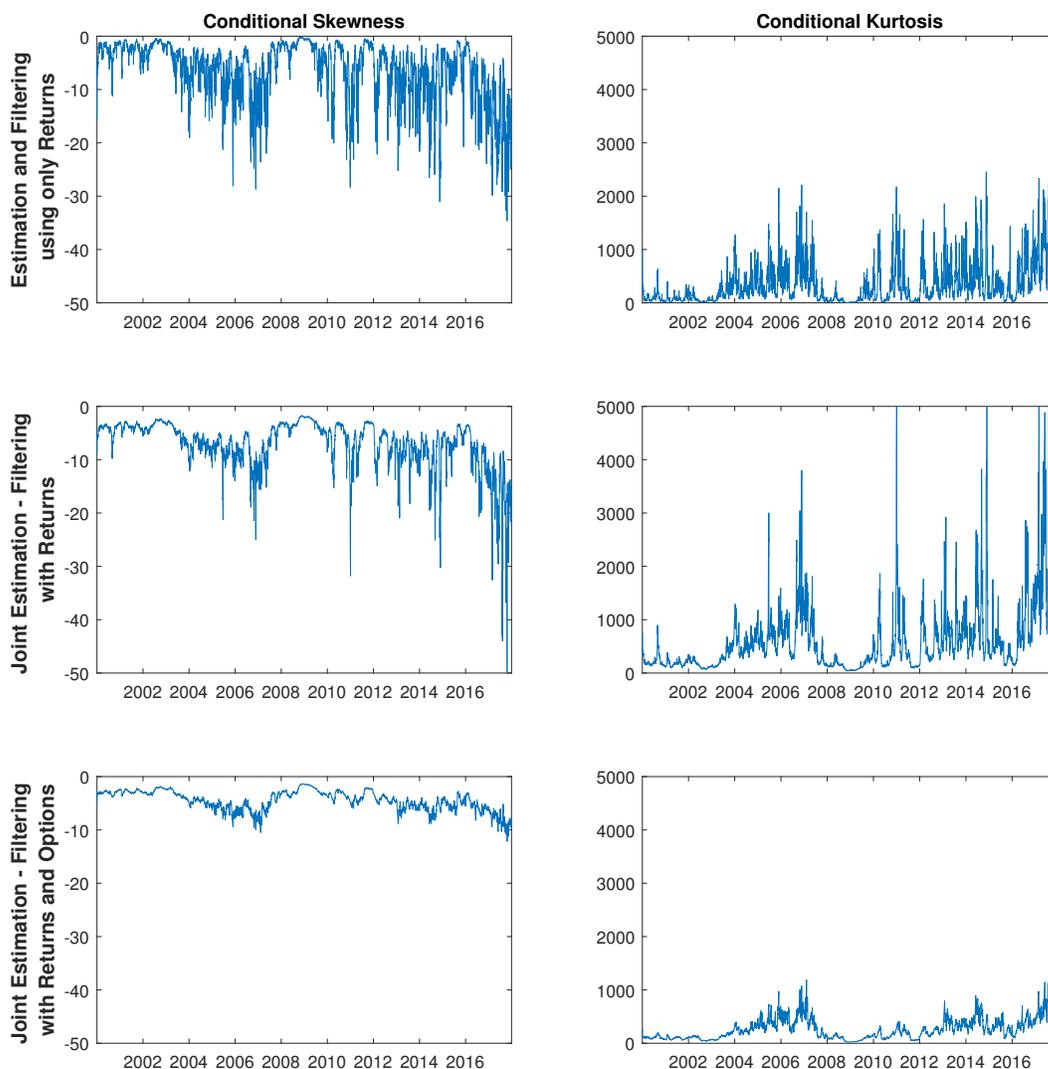
From the filtered conditional variance and jump intensity, we compute the contribution of each risk source to the equity risk premium ($252 \times (\lambda_z h_z + \lambda_y h_y)$) of the S&P500 daily returns. The first plot report the filtered equity risk premium when only returns are used in the parameter estimation, the second one when latent states are filtered with returns and parameter estimation is done with a two stages weighted likelihood, and the third one when returns and options are used to filter latent states.

Figure 9: Conditional volatility



From the filtered conditional variance and jump intensity, we compute the annualized conditional volatility ($\sqrt{252 \times (h_{z,t+1} + \psi''_{y_{t+1}|\mathcal{F}_t}(0))}$) of the S&P500 daily returns. See A.2 for details. The first plot report the filtered conditional volatility when only returns are used in the parameter estimation, the second one when latent states are filtered with returns and parameter estimation is done with a two stages weighted likelihood, and the third one when returns and options are used to filter latent states.

Figure 10: Higher order conditional moments



From the filtered conditional variance and jump intensity, we compute higher order conditional moments of the S&P500 daily returns. See [A.2](#) for details. The top two plots report the filtered conditional skewness and kurtosis when only returns are used in the parameter estimation, the middle two when latent states are filtered with returns and parameter estimation is done with a two stages weighted likelihood, and the bottom two when returns and options are used to filter latent states.

Table 9: Higher order conditional moments

	Estimation and Filtering using only Returns	Joint Estimation - Filtering with Returns	Joint Estimation - Filtering with Returns and Options
Percentage of variance			
From diffusion	71.14%	78.50%	78.93%
From jumps	28.86%	21.50%	21.07%
Avg. cond. volatility	15.79%	17.22%	16.07%
Avg. cond. skewness	-6.41	-7.07	-4.50
Avg. cond. kurtosis	344.12	613.01	265.11

Average filtered conditional moments of the daily S&P500 index's returns. Volatility is annualized.

Table 11: Summary of the sample of options used for model estimation

Panel A: Number of option contracts			
DTM]10,45]]45,90]	All
Moneyiness			
]0.775,0.850]	3619	4552	8171
]0.850,0.925]	5365	5373	10738
]0.925,1.000]	5449	5503	10952
]1.000,1.075]	5451	5456	10907
]1.075,1.150]	2416	4002	6418
All	22300	24886	47186
Panel B: Average implied volatility			
DTM]10,45]]45,90]	All
Moneyiness			
]0.775,0.850]	0.3657	0.3067	0.3328
]0.850,0.925]	0.2806	0.2525	0.2666
]0.925,1.000]	0.2086	0.2033	0.2060
]1.000,1.075]	0.1673	0.1703	0.1688
]1.075,1.150]	0.2000	0.1640	0.1775
All	0.2404	0.2193	0.2293

Panel A summarize the number of contracts for each moneyiness and maturity buckets and Panel B their average implied volatility.