Risk Premium Inference from (Noisy) Option Price Panels

First version: June 16, 2018

Previous Version : February 21, 2024

This version: March 28, 2024

Brief Description of Research Idea

Compare the finite sample performance of implied state vs. particle filtering when conducting joint $\mathbb{P}$-$\mathbb{Q}$ estimation on noisy option price panels focusing on the implications this has for recovered (equity) risk premium components and their error distributions/confidence bands.

Tentative Outline of Building Blocks

1. Implement an efficient simulation scheme for a Heston (Heston, 1993) state space simulation and associated option price panels.

2. Implement IS estimation à la Boswijk et al. (2015).

3. Implement particle filtering à la Hurn et al. (2015).

4. No noise/best case.

5. Noisy option panel case with an assorted menu of assumptions regarding option price error distributions.

Rationale

A. Noisy option data is something I had to deal with both for the first two papers, outside academia and having spent some more time looking at post 2018 academic work it seems that this is an emerging topic with a few papers that touch on this in a different context: Andersen et al. (2021), Duarte et al. (2019).

B. Topic itself ties in well with the other two papers in my thesis.

C. I have sufficient background to work on this independently.

1 Assumed Data Generating Proces And Option Pricing

We consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions[1]. Under the historical probability measure, $\mathbb{P}$, the dynamics of the log-price process $\log(S_t)$ and of the stochastic volatility process V_t, which make up the state vector $X_t = [\log(S_t), V_t]$, are assumed to be described by the pair of s.d.e.:

$$\mathrm{dlog}(S_t) = \left((r_t - q_t) + (\eta - 1/2)\,V_t\right)\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_{t,1}^{\mathbb{P}}; \tag{1.1}$$

$$\mathrm{d}V_t = k(\bar{V} - V_t)\,\mathrm{d}t + \sigma_v\sqrt{V_t}\left(\rho\,\mathrm{d}W_{t,1}^{\mathbb{P}} + \sqrt{1 - \rho^2}\,\mathrm{d}W_{t,2}^{\mathbb{P}}\right). \tag{1.2}$$

$\left(\mathrm{d}W_{t,1}^{\mathbb{P}}, \mathrm{d}W_{t,2}^{\mathbb{P}}\right)$ are $\mathcal{F}_t$-adapted Brownian motion processes. $\{r\}_{t\geq 0}$ and $\{q\}_{t\geq 0}$ are $\mathcal{F}_t$-adapted processes (e.g., deterministic) which describe the risk-free rate of interest and the continuous dividend yield respectively. η is a parameter which governs the size of the equity risk premium. The local stochastic volatility process mean reverts at a mean reversion rate κ to the long run mean level $\bar{V}$. σ_v governs the volatility of volatility and ρ represents the instantaneous correlation between returns and volatility.

Under the risk neutral probability measure, $\mathbb{Q}$, the model dynamics are assumed to be:

$$\mathrm{dlog}(S_t) = \left((r_t - q_t) - 1/2V_t\right)\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_{t,1}^{\mathbb{Q}}; \tag{1.3}$$

$$\mathrm{d}V_t = k(\bar{V} - V_t)\,\mathrm{d}t + \eta_V V_t\,\mathrm{d}t + \sigma_v\sqrt{V_t}\left(\rho\,\mathrm{d}W_{t,1}^{\mathbb{Q}} + \sqrt{1 - \rho^2}\,\mathrm{d}W_{t,2}^{\mathbb{Q}}\right). \tag{1.4}$$

$\left(\mathrm{d}W_{t,1}^{\mathbb{Q}}, \mathrm{d}W_{t,2}^{\mathbb{Q}}\right)$ are Brownian motions under the (equivalent) martingale pricing measure $\mathbb{Q}$.

Jointly, 1.3 and 1.4 are known as the Heston model, which features local stochastic volatility with a leverage effect while providing semi-closed form solutions for option prices. This model is a "staple" stochastic volatility model for equity index modeling. It is used in a large number of empirical studies, has a number of well documented limitations, some of which have been remedied by enriching the model specification through the addition of jumps and/or other (latent) stochastic states.

Given the focus of the study, attention is restricted to the simplistic set-up of the Heston model for a host of reasons. On the one hand, for most Heston model extensions, econometric approaches to parameter estimation which employ option price data come with increased computational costs. Since in the current set-up this model is known

---

[1]See e.g., Protter (2005), p. 3.

to be the data generating process, there's no potential misspecification issue at play. Secondly, extensions come with specific estimation challenges which typically require tailored estimation designs to obtain well-behaved parameter estimates therefore limiting the comparability of results obtained throughout the study.

1.1 Parameter Sets

Let θ denote the vector of parameters featured in the model. We assume $\theta \in \Theta$ where Θ is a compact set from the domain of valid data-generating process parameters. For the Heston model specification previously introduced, the value of θ corresponding to the data-generating process introduced in (1.1) - (1.4) is $\theta = [\eta, \kappa, \bar{v}, \sigma_v, \rho, \eta_v]$. Note that subsets of the parameter set play a role in either the $\mathbb{P}$-measure specification, the $\mathbb{Q}$-measure specification or in both. E.g., the Heston parameter set can be decomposed into 3 disjoint subsets: $\theta = \left[\theta^{\mathbb{P}}, \theta^{\mathbb{P} \cap \mathbb{Q}}, \theta^{\mathbb{Q}}\right]$, where $\theta^{\mathbb{P}} = \eta$, $\theta^{\mathbb{P} \cap \mathbb{Q}} = [\kappa, \bar{v}, \sigma_v, \rho]$, and $\theta^{\mathbb{Q}} = \eta_v$.

1.2 State Vector

Assuming we have a time series of state vector observations:[2] $X_1, \ldots, X_T$, observed at equally-spaced timepoints $t_1, \ldots, t_T$ with $t_{i+1} - t_i = \Delta > 0, \forall i \geq 1$. Denote by $f(X_{i+1}|X_i; \theta^{\mathbb{P}}, \theta^{\mathbb{P} \cap \mathbb{Q}})$ the transition density of the state vector X_t under the historical probability measure. Note that the transition density depends on $\theta^{\mathbb{P}}$ and $\theta^{\mathbb{P} \cap \mathbb{Q}}$ which represent the subset of parameters that describe the model dynamics under $\mathbb{P}$.

By using option price data the full set of parameters – including the additional ones that characterize model dynamics under the risk neutral pricing measure – can be estimated, i.e., inference on $\theta^{\mathbb{Q}}$ can be conducted while at the same time increasing sample identification potential for the $\theta^{\mathbb{P} \cap \mathbb{Q}}$ parameter set.

1.3 Option prices

Suppose that, in addition to time series observations of returns, synchronous observations of option price panel data and time series data on the underlying are available. Denote the option prices by $p_i(T, K)$ where T, K represent the maturity and the strike of the option contract. Option prices are commonly quoted in terms of their Black-Scholes implied volatility. Let $P(X_i, T, K; \theta^{\mathbb{P} \cap \mathbb{Q}}; \theta^{\mathbb{Q}})$ be the corresponding model price, e.g., the option price generated by the base model specification under the equivalent martingale

---

[2]In practice such a series never exists as the volatility, V_t, is an unobservable, i.e., latent state process

pricing measure for given state vector instance X_i and parameter set values $\theta^{\mathbb{P}\cap\mathbb{Q}}$ and $\theta^{\mathbb{Q}}$. If option prices were observed without error and the model is not misspecified, the observed option prices would coincide with their model-based counterparts. In practice the two prices differ. Typical modelling approaches assume away model misspecification and attribute the difference to measurement errors (e.g., microstructure noise).

1.3.1 Particle Filter Based Estimation

Estimation contexts which make use of option prices (that are nonlinear, often semi-analytic expressions featuring model parameters and partially observable state vectors) can be computationally demanding. The models used to describe index returns for option pricing purposes typically feature latent stochastic state processes which require filtering or implying. This increases the computational cost of econometric estimation exercises as one has to devise parsimonious estimators which can work with non-analytic expressions of state vector observations by relying on stable and robust numerical procedures.

Two other hurdles are typically encountered when working with continuous time models. Firstly, even the simple continuous time stochastic model specifications (like the base-case model used in this paper) do not have known densities readily usable to evaluate the likelihood function. Secondly, one often has to resort to using discretisation schemes to simulate state vector (i.e., particle) paths. However, despite this second hurdle, approximating transition densities using discrete distributions consisting of particle support points provides a tractable workaround which can easily be adapted for most of the typical model specifications considered. The filtering distributions provide estimators for the latent states and their possible distributions. The Monte Carlo procedures involved in the simulation of particles can be tailored to accommodate models with jumps and with more complex state dynamics than featured in the Heston model specification.

In the context of the base model specification previously introduced, we outline the following optimal filtering procedure in the spirit of Johannes et al. (2009), Christoffersen et al. (2010) and ***[Hurn et al]*** and introduce the necessary implementation details for the Monte Carlo estimation exercise. First, let us denote the log-likelihood function of the model under the physical probability measure by

$$\log \mathcal{L}_T = \log f(X_0) + \sum_{i=0}^{T} \log f(X_{i+1}|X_i; \theta^{\mathbb{P},\mathbb{P}\cap\mathbb{Q}}). \qquad (1.5)$$

The evaluation of this likelihood function is unfeasible as the state vector is not fully

observed. Also note that it does not provide any information about $\theta^{\mathbb{Q}}$, the parameters which feature in the dynamics of the model under the risk neutral probability measure.

To take into account option prices and to allow for the likelihood to be evaluated, a filtering rule using the time series observations for the components of the state vector which are observable is introduced alongside an option price panel likelihood observation in order to extract information about the latent states which are not observable. Let us denote by $\mathcal{O}_i$ the filtration containing the collection of sigma algebras generated by the time series of observable states and associated option prices. E.g., for our base model specification this is the filtration generated by $S_1, \ldots, S_T$ and the panel $p_1(T, K), \ldots p_T(T, K)$ of option prices for all available option maturities T and strike prices K. The level of the stochastic volatility process V_i is unobservable. To evaluate the likelihood contribution, the following need to be evaluated:

Step 1. T_{i+1} likelihood contribution:

$$
\begin{aligned}
f(X_{i+1}|\mathcal{O}_i; \theta) &= \int_{\mathcal{D}_V} f(S_{i+1}|V_{i+1}, \mathcal{O}_i; \theta) f(V_{i+1}|\mathcal{O}_i; \theta)\, \mathrm{d}V_{i+1} \\
&= \int_{\mathcal{D}_V} f(S_{i+1}|V_{i+1}, \mathcal{O}_i; \theta) \int_{\mathcal{D}_V} f(V_{i+1}|V_i, \mathcal{O}_i; \theta) \\
&\qquad\qquad\qquad\qquad\qquad \times f(V_i|\mathcal{O}_i; \theta)\, \mathrm{d}V_i\, \mathrm{d}V_{i+1}
\end{aligned}
\tag{1.6}
$$

Step 2. Latent state variable density update:

$$
f(V_{i+1}|\mathcal{O}_{i+1}; \theta) = \frac{f(X_{i+1}, V_{i+1}|\mathcal{O}_{i+1}; \theta)}{f(X_{i+1}|\mathcal{O}_{i+1}; \theta)}
\tag{1.7}
$$

References

Ait-Sahalia, Y. and Kimmel, R. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of financial economics*, 83(2):413–452.

Andersen, T., Fusari, N., Todorov, V., and Varneskov, R. (2021). Spatial dependence in option observation errors. *Econometric Theory*, 37(2):205–247.

Bardgett, C., Gourier, E., and Leippold, M. (2014). Inferring volatility dynamics and risk premia from the s&p 500 and vix markets.

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

Boswijk, H. P., Laeven, R. J., and Lalu, A. (2015). Asset returns with self-exciting jumps: Option pricing and estimation with a continuum of moments. *Working Paper*.

Christoffersen, P., Jacobs, K., and Mimouni, K. (2010). Volatility dynamics for the s&p500: evidence from realized volatility, daily returns, and option prices. *The Review of Financial Studies*, 23(8):3141–3189.

Duarte, J., Jones, C. S., and Wang, J. L. (2019). Very noisy option prices and inference regarding the volatility risk premium. *Available at SSRN 4167308, forthcoming Journal of Finance*.

Garcia, R., Lewis, M.-A., Pastorello, S., and Renault, É. (2011). Estimation of objective and risk-neutral distributions based on moments of integrated volatility. *Journal of Econometrics*, 160(1):22–32.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Hurn, A. S., Lindsay, K. A., and McClelland, A. J. (2015). Estimating the parameters of stochastic volatility models using option price data. *Journal of Business & Economic Statistics*, 33(4):579–594.

Johannes, M. S., Polson, N. G., and Stroud, J. R. (2009). Optimal filtering of jump diffusions: Extracting latent states from asset prices. *The Review of Financial Studies*, 22(7):2759–2799.

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

Protter, P. E. (2005). *Stochastic integration and differential equations.* Springer.

Sundaresan, S. M. (2000). Continuous-time methods in finance: A review and an assessment. *The Journal of Finance*, 55(4):1569–1622.