Joint Calibration Thesis

Mustafif Khan

Abstract Coming Soon!

September 24, 2024

Contents

1. Introduction

The financial modelling world tends to look at American Options with a shy eye, to be able to be exercised at any day before the maturity presents unique challeges compared to simpler styles like the European Options. When we take a look at timeseries way to forecast American options pricing, we often either consider the *daily log returns* of the underlying asset **or** *implicit volatility* of the option prices. So we have estimation taking into account only one measure, we are either looking at historical data of the underlying asset in the Physical measure or looking forward with the implicit volatility which is in the Risk-Neutral measure.

Definition: Risk-Neutral Measure

A probability measure such that the expectation of the returns for all assets is equivalent to the risk-free rate. Under this paper, we will denote this measure as $\mathbb{Q}$.

Definition: Physical Measure

A probability measure that takes into account the real-world proabilities that reflect the actual likelihood of events which occur in the financial markets. Under this paper, we will denote this measure as $\mathbb{P}$.

Our goal in this paper is to take into account both of these probability measures to better estimate American Options under a GARCH model. To accomplish this, we will be using a technique called **Joint Calibration** which will be used as an objective function for an Artificial Neural Network which will use both historical asset prices and implicit volatilities to calibrate GARCH model parameters.

1.1. Motivation

TODO

Why didnt traditional optimization work?

- Willow Tree not being efficient, very slow with this idea
- What exactly were the problems with the initial idea of this
- What Artificial Neural Networks comes to try and solve, with the advantages it brings

1.2. Related Works

“Estimating and using GARCH models with VIX data for option valuation” by Juho Kanniainen, Binghuan Lin and Hanxue Yang.

- This paper utilizes the GARCH model and the Joint MLE to use information on VIX to improve the empirical performance of the models for pricing options on the S&P 500.
- VIX (Cboe Volatility Index) follows a European Exercise Style, our paper focuses on American Options which will be harder to price, and take into account the more complex exercise style.

“FX Volatility Calibration Using Artificial Neural Networks” by Alexander Winter, Kellogg College, University of Oxford

- This paper explores effectively learning model parameters using an ANN with the Heston stochastic volatility model, and how it can perform orders of magnitudes faster than traditional optimization-based methods.
- This paper is another great example of using ANN as an application to calibrate financial models, which we also hope to demonstrate in this paper, while this paper focuses on the Heston Model, and we focus on GARCH model, we hope to apply some of the ideas brought in this paper like data gathering to aid us.

1.3. Structure

TODO

- This is where we will discuss why we will use two different GARCH models
- An initial look at a simple Artificial Neural Network
- Use some of the summary parts for this
- What American Option we will use, specifications
- Training idea, synthetic data could come here or a “Data Gather” section

2. Estimation with GARCH Models

In our paper we will be taking a look at two different GARCH models and how they are represented under both $\mathbb{P}$ and $\mathbb{Q}$ measures. The two models we will take a closer look is the Heston-Nandi GARCH(1, 1) model and the Duan (1995) model.

2.1. GARCH Processes

Before we are able to look, we need the general picture of the GARCH process where under the physical measure $\mathbb{P}$, we assume the underlying stock price process follows a conditional distribution D where we can express the log returns R_t at time period t as:

$$\begin{aligned} R_t \equiv \ln\left(\frac{S_t}{S_{t-1}}\right) &= \mu_t - \gamma_t + \varepsilon_t \qquad && \varepsilon_t \mid F_{t-1} \sim D(0, h_t) \\ &= \mu_t - \gamma_t + \sqrt{h_t} z_t && z_t \mid F_{t-1} \sim D(0, 1) \end{aligned} \tag{1}$$

$\varepsilon_t = \sqrt{h_t} z_t$

- S_t: Stock price at time t
- h_t: Conditional variance of the log return in period t

2.2. Parameter Definitions

- ω: Long term average variance constant
- α: Coefficient for lagged innovation
- β: Coefficient for lagged variance
- γ: The Asymmetry coefficient
- λ: Price of risk or risk premium
- r: Risk-free rate or discount rate

Relationships:

- α and β ensures model's stationarity ($\alpha + \beta < 1$)
- $\gamma > 0$ indicates negative shocks have a larger impact on future volatility.

2.3. Risk Neutralization

- $\mathbb{P}$: Physical Measure
- $\mathbb{Q}$: Risk-Neutral Measure

Using the Radon-Nikodym derivative, we can convert the physical measure to the risk-neutral measure. Let z_t i.i.d $N(0, 1)$, then $\gamma_t = \frac{1}{2} h_t$ since $\exp(\gamma_t) = E_{t-1}[\exp(\varepsilon_t)]$

The Radon-Nikodym derivative is defined as:

$$\frac{\mathrm{dQ}}{\mathrm{dP}} \mid F_t = \exp\left(-\sum_{i=1}^{t}\left(\frac{\mu_i - r_i}{h_i}\varepsilon_i + \frac{1}{2}\left(\frac{\mu_i - r_i}{h_i}\right)^2 h_i\right)\right) \tag{2}$$

We also get under the RN-Derivative:

Defined by $\varepsilon_t \mid F_{t-1} \sim N(-(\mu_t - r_t), h_t)$

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r_t - \frac{1}{2}h_t + \varepsilon_t^* \tag{3}$$

with $\varepsilon_t^* \mid F_{t-1} \sim N(0, h_t)$ and $E^Q\left[\frac{S_t}{S_{t-1}} \mid F_{t-1}\right] = \exp(r_t)$

2.3.1. Heston-Nandi GARCH

$\mathbb{P}$:

$$\begin{aligned} R_t \equiv \ln\left(\frac{S_t}{S_{t-1}}\right) &= r + \lambda h_t + \varepsilon_t \\ h_t &= \omega + \beta h_{t-1} + \alpha\left(\varepsilon_{t-1} - c\sqrt{h_{t-1}}\right)^2 \end{aligned} \tag{4}$$

Assume:
- $r_t = r$
- $\mu_t = r + \left(\lambda + \frac{1}{2}\right)h_t$

To risk neutralize Equation 4 we subsitute it along with the assumptions stated into Equation 2 to get the corresponding RN-Derivative for the Heston-Nandi Model:

$$\frac{\mathrm{dQ}}{\mathrm{dP}} \mid F_t = \exp\left(-\sum_{i=1}^{t}\left(\left(\lambda + \frac{1}{2}\right)\varepsilon_i + \frac{1}{2}\left(\lambda + \frac{1}{2}\right)^2 h_i\right)\right) \tag{5}$$

Risk-neutral innovations of the form:

$\varepsilon_t^* = \varepsilon_t + \lambda h_t + \frac{1}{2}h_t$

$\mathbb{Q}$:

$$\begin{aligned} R_t \equiv \ln\left(\frac{S_t}{S_{t-1}}\right) &= r - \frac{1}{2}h_t + \varepsilon_t^* \\ h_t &= \omega + \beta h_{t-1} + \alpha\left(z_{t-1}^* - \left(\gamma + \lambda + \frac{1}{2}\right)\sqrt{h_{t-1}}\right)^2 \end{aligned} \tag{6}$$

Using $z_t^* \overset{Q}{\sim} N(0,1)$, $\varepsilon_t^* = \sqrt{h_t}z_t^*$ and $\rho^* = \gamma + \lambda + \frac{1}{2}$ into Equation 6 we get Equation 7:

$$\begin{aligned} R_t &= \left(r - \frac{1}{2}h_t\right) + \sqrt{h_t}z_t^* \\ h_t &= \omega + \beta h_{t-1} + \alpha\left(z_{t-1}^* - \rho^*\sqrt{h_{t-1}}\right)^2 \end{aligned} \tag{7}$$

2.3.2. Duan (1995)

$\mathbb{P}$:

$$
\begin{gathered}
R_t \equiv \ln\left(\frac{S_t}{S_{t-1}}\right) = r_t + \lambda\sqrt{h_t} - \frac{1}{2}h_t + \varepsilon_t \\
h_t = \omega + \beta h_{t-1} + \alpha\varepsilon_{t-1}^2
\end{gathered} \tag{8}
$$

Assume:

- Price of risk λ is assumed to be constant
- $r_t = r$
- $\mu_t = r + \lambda\sqrt{h_t}$ or $\lambda = \frac{\mu_t - r}{\sqrt{h_t}}$

To risk neutralize Equation 8 we subsitute it along with the assumptions stated into Equation 2 to get the corresponding RN-Derivative for the Duan (1995) Model:

$$
\frac{\mathrm{dQ}}{\mathrm{dP}} \mid F_t = \exp\left(-\sum_{i=1}^{t}\left(\frac{\varepsilon_i}{\sqrt{h_i}}\lambda + \frac{1}{2}\lambda^2\right)\right) \tag{9}
$$

Risk-neutral innovations of the form:

$\varepsilon_t^* = \varepsilon_t + \mu_t - r_t = \varepsilon_t + \lambda\sqrt{h_t}$

$\mathbb{Q}$:

$$
\begin{gathered}
R_t \equiv \ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}h_t + \varepsilon_t^* \\
h_t = \omega + \beta h_{t-1} + \alpha\left(\varepsilon_{t-1}^* - \lambda\sqrt{h_{t-1}}\right)^2
\end{gathered} \tag{10}
$$

Let $\varepsilon_t^* = z_t^*\sqrt{h_t}$ with $z_t^* \overset{Q}{\sim} N(0,1)$:

$$
\begin{gathered}
R_t = r - \frac{1}{2}h_t + z_t^*\sqrt{h_t} \\
h_t = \omega + \beta h_{t-1} + \alpha\left(\sqrt{h_{t-1}}(z_{t-1}^* - \lambda)\right)
\end{gathered} \tag{11}
$$

2.4. Log Likelihoods

2.4.1. Return Log Likelihood

The log likelihood of the return process is calculated under the $\mathbb{P}$ measure.

Let Y_1 represent the returns log likelihood, h_i be the daily returns, and N be the number of days in the returns sample, we can then compute it as:

$$Y_1 = -\frac{1}{2}\sum_{i=1}^{N}\left\{\ln(h_i) + \frac{(R_i - \mu_i + \gamma)^2}{h_i}\right\} \tag{12}$$

From the GARCH Process, we can notice that $R_i - \mu_i + \gamma$ in the formula is equivalent to the following:

$$\begin{aligned} R_t &= \mu_t - \gamma + \sqrt{h_t} z_t \\ \sqrt{h_t} z_t &= R_t - \mu_t + \gamma \end{aligned} \tag{13}$$

Substituting the above into Equation 12 we get the following:

$$Y_1 = -\frac{1}{2}\sum_{i=1}^{N}\{\ln(h_i) + z_i^2\} \qquad z_i \sim D(0,1) \tag{14}$$

NOTE:

This derivation with the formula with Equation 14, allows us to calculate the Log Likelihood agnostic towards whichever GARCH model we'd like to use, whether Heston-Nandi or Duan.

2.4.2. Options Log Likelihood

The log likelihood of the options process is calculated under the $\mathbb{Q}$ measure.

Assume

$$\underbrace{\sigma_{i,t}}_{\text{Imp vol. from market price}} = \underbrace{\sigma_{i,t}(C_{i,t}(h_t(\xi^*)))}_{\text{Imp vol. from GARCH model}} + \varepsilon_{i,t} \tag{15}$$

where $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$

Let Y_2 represent the options log likelihood and M be the number of options, we can then compute it as:

$$Y_2 = -\frac{1}{2}\sum_{i=1}^{M}\left\{2\ln(\sigma_\varepsilon) + \frac{\left(\overbrace{\sigma_i}^{\text{Imp vol. on market price}} - \overbrace{\sigma_{\text{imp},i}}^{\text{Impl vol. from models}}\right)^2}{\sigma_\varepsilon^2}\right\} \tag{16}$$

2.4.3. Joint Log Likelihood

$$Y_{\text{joint}} = \frac{N+M}{2N} Y_1 + \frac{N+M}{2M} Y_2 \tag{17}$$

In our Calibration Artificial Neural Network, we will use $-Y_{\text{joint}}$ as our objective function to to solve for the parameters $\theta = (\omega, \alpha, \beta, \gamma, \lambda, \sigma_\varepsilon)$, with the non-negativity constraint on the parameters in θ.

3. Conclusion

Bibliography