Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrie & Operationele Research

A Simple HAR Option Pricing Framework

Casper Keer (613906)

Supervisor:	dr. P. Vallarino
Second assessor:	dr. G. Freire
Date final version:	1st July 2024

Abstract

This thesis proposes a simple option pricing framework based on the Heterogeneous Autoregressive (HAR) model and Monte Carlo simulations. Forecasts of realized volatility, a nonparametric measure for asset variation, are made with the HAR model, which are then used in an Euler scheme for asset path generation. To address pseudo-constant volatility due to converging predictions, an extension that includes random jumps is presented. Particular attention is paid to the fastest-growing and least literature-covered maturity spectrum: weekly and biweekly options. In an empirical study on out-of-the-money S&P 500 index options, the proposed framework is found to outperform Black-Scholes and GARCH benchmarks for non-extreme moneyness categories. Finally, robustness is assessed by considering a sample period of extreme events. Although performance deteriorates, high pricing accuracy is found for call options.

1 Introduction

Volatility stands as a cornerstone in financial markets, serving as a decisive component for traders and investors alike. Effective risk management and accurate valuation of derivatives hinge upon its precise measurement. Realized volatility is a nonparametric ex-post measure of asset variation constructed from the sum of intraday squared returns. Due to its favorable statistical properties[1] and ability to react quickly to asset price movements, it leads to accurate forecasts of future asset variation when proper modeling is done. One of the most widely used models to forecast this measure is the Heterogeneous Autoregressive model for realized volatility (HAR-RV) of Corsi (2009). This parsimonious model originates from the Heterogeneous Market Hypothesis introduced by Müller et al. (2008), which acknowledges the existence of diversity among traders in the market. Market participants have different time horizons and trading frequencies: market makers tend to have all positions closed by the market close, while central banks and hedge funds can stay in one position for weeks or even months. Hence, different trading frequencies imply different reactions to the same news in the same market. Inspired by this phenomenon and the promising literature on applications of the HAR-RV model in option pricing (see, e.g., Corsi, Fusari and La Vecchia (2013) and Alitab, Bormetti, Corsi and Majewski (2019)), this thesis presents a straightforward option pricing framework based on the HAR-RV model. In contrast to previous literature, however, particular attention is given to short-maturity (i.e., below one month) options, which are commonly omitted. This is somewhat surprising as the market for so-called "weeklies", options expiring one week apart, has seen a sharp increase since 2012 and became the most traded segment of the maturity spectrum by the middle of 2015 (Andersen, Fusari & Todorov, 2017).

Apart from this enlarging option market, the need for a new option pricing framework lies in the simplicity argument. HAR option valuation models, as in Corsi et al. (2013), assume a conditional distribution for RV, from which the parameters must be obtained each time period by Maximum Likelihood Estimation (MLE). This framework is particularly useful if the number of options to be priced is relatively small, and parameters should not be estimated frequently. Indeed, the authors of the previously cited paper filter out options with maturities below nine

[1] For increasing sampling frequency (and absence of microstructure noise), realized volatility converges to the quadratic variation of the asset return process (see, e.g., Barndorff-Nielsen and Shephard (2002)).

days and mainly observe options with maturities above 20 days. This is understandable, as Corsi et al. (2013) (and papers building upon it, such as Alitab et al. (2019)) consider a sample prior to the increase in popularity of weeklies. Furthermore, the common absence of model performance analysis in periods of turmoil (e.g., the Financial Crisis of 2008 and the COVID-19 Pandemic of 2020) is asking for a more profound investigation on the robustness of option pricing models based on realized volatility.

The contribution of this thesis is threefold. First, an option pricing framework based on the HAR-RV model of Corsi (2009) is presented, and an extension is provided which incorporates jumps, one of the most characteristic features of asset returns. The framework is used to price out-of-the-money (OTM) S&P 500 index options with maturities ranging from two days to six months. For options that are not deep-OTM, average improvements of around 20% and 6% in terms of Root Mean Squared Error on Implied Volatility (RMSE_{IV}) are obtained relative to the Practitioners Black-Scholes model of Dumas, Fleming and Whaley (1998) and GARCH option pricing model of Heston and Nandi (2000), respectively. Second, particular attention is paid to the commonly omitted short-maturity options (maturities between two and twenty days). For this maturity spectrum, the proposed option pricing framework yields the most accurate prices for non-extreme moneyness options and call options. Third, a separate analysis is done on pricing performance in periods of economic turmoil. It is found that the framework achieves higher RMSE_{IV} across moneyness and maturities, except for deep-OTM calls. This shows that the proposed framework captures the right tail of the volatility distribution quite well in periods of high volatility.

The remainder of this paper is structured as follows. Section 2 reviews related literature on volatility modeling and option pricing. In Section 3, a theoretical foundation and empirical application of the HAR-RV model are provided. Section 4 describes the option pricing framework, after which Section 5 applies the proposed methods in an empirical setting. Section 6 concludes this thesis.

2 Related Literature

Financial returns possess several well-known stylized facts, including slowly decreasing autocorrelation of squared and absolute returns, fat-tailed distribution with shapes depending on the frequency (low-frequency returns approach a normal distribution), and no significant autocorrelation in returns. Several models have been proposed that try to replicate these properties. The most widely known being the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of Bollerslev (1986), and Stochastic Volatility (SV) models (see, e.g., Melino and Turnbull (1990) and Jacquier, Polson and Rossi (2002)). However, these models present two issues: they fail to capture the slowly decaying autocorrelation in squared returns and are not straightforward to estimate. Therefore, a growing interest emerged in long-memory volatility models. Inspired by Granger and Joyeux (1980), the fractionally integrated model for realized volatility was introduced by Andersen, Bollerslev, Diebold and Labys (2003). It incorporates long-memory in a parsimonious way by using fractional integration. This model, however, lacks interpretation (it is unclear what the economic meaning of fractional integration is) and is tedious to implement. Corsi (2009) addresses these issues by presenting the previously

mentioned HAR-RV model, which, despite being simple and not formally belonging to the class of long-memory models, is able to reproduce the main stylized facts of financial returns.

Option pricing has caught the most steam due to the introduction of the Black and Scholes (1973) option pricing model (henceforth BS). It takes the main determinants of option prices (underlying asset price, strike price, volatility of the underlying, time to maturity, dividend yield, and interest rates) as inputs and outputs the corresponding option price. However, the BS model is nowadays rarely used for the pricing of options, and more for the derivation of the implied volatility (i.e., the market-expected volatility).[2] This is because the BS model assumes constant volatility of the underlying over the option's lifetime and hence fails to exhibit the so-called "volatility smile": different implied volatilities for different strike prices at a fixed maturity.

From the constant volatility framework of BS, there are two categories of extensions: discrete-time GARCH models and continuous-time SV models. The GARCH option pricing models are based on the assumption that volatility depends on its lagged values and lagged squared returns (see, e.g., Heston and Nandi (2000) and Christoffersen, Jacobs, Ornthanalai and Wang (2008)). In the SV framework, time is considered to be continuous, and a stochastic process is assumed for volatility (see, e.g., Hull and White (1987) and Heston (1993)). In both categories of models, however, volatility is considered unobserved, which complicates their application in empirical option pricing.

With the availability of high-frequency data, one can now use realized volatility as a proxy for the "true" asset variation (i.e., the quadratic variation of the asset return process). As a result, Christoffersen, Feunou, Jacobs and Meddahi (2014) extend the GARCH option pricing model of Heston and Nandi (2000) by including realized volatility measures, and Corsi et al. (2013) construct a HAR option pricing model. The general idea in both papers is to assume processes for asset returns and realized volatility. For example, Corsi et al. (2013) assume dynamics for logarithmic returns that are conditional on realized volatility, which, in turn, follows an autoregressive gamma process. After parameter estimation, option prices are obtained using Monte Carlo simulations.[3] This method of pricing options, advocated by Boyle, Broadie and Glasserman (1997), consists of simulating many asset paths, after which the value of an option is calculated as the discounted average option payoff of the paths at expiration. Due to its simplicity, the framework proposed by this thesis also uses simulation methods for option pricing.

3 Preliminary Analysis

In this section, details on realized volatility are discussed and the HAR-RV model is introduced. Furthermore, the model's ability to predict realized volatility is evaluated by means of a forecasting exercise on a major S&P 500 ETF, comparing its accuracy over different forecasting horizons against several benchmarks.[4]

[2]The option price is a known variable from the market's supply and demand forces. Hence, the BS equation can be numerically solved to obtain the market implied volatility since all other inputs are known.

[3]The model of Corsi et al. (2013) allows for a semi-closed-form expression of the option price, which is not used for longer maturities due to computational costs. Hence, Monte Carlo simulations were used for this spectrum.

[4]This section replicates Corsi (2009), but with a different dataset and additional hypothesis tests.

3.1 The HAR-RV Model

As is common in literature, the logarithmic asset price $X(t) = \log S(t)$ is assumed to evolve according to a diffusion process with Stochastic Differential Equation (SDE) of the form

$$dX(t) = \mu(t)dt + \sigma(t)dW(t), \quad 0 \leq t \leq T, \tag{1}$$

where $\mu(t)$ is a càdlàg[5] finite variation process, $\sigma(t)$ is a strictly positive stochastic process, and $W(t)$ is a standard Brownian motion. In this setting, the quadratic variation of the cumulative return process over a one-day interval is defined as

$$QV_t = \int_{t-1}^{t} \sigma^2(s)\, ds, \tag{2}$$

which is also referred to as the integrated variance.[6] In the remainder of this thesis, the subscript t refers to day t, but in reality, could be any unit of time. In the absence of microstructure noise, the daily quadratic variation can be approximated arbitrarily well using the sum of M intraday squared returns. Define the daily realized volatility as

$$RV_t = \sqrt{\sum_{i=1}^{M} r_{t_i}^2}, \tag{3}$$

where $r_{t_i} = X(t-1+i/M) - X(t-1+(i-1)/M)$ is the continuously compounded intraday return sampled at time interval $1/M$. For an increasing sampling frequency (i.e., $M \to \infty$), it holds that $RV_t^2 \xrightarrow{p} QV_t$, where $\xrightarrow{p}$ stands for convergence in probability. However, high-frequency data generally suffers from microstructure noise such as price discreteness and rounding, bid-ask bounces, and differences in information. Therefore, as in Corsi (2009), this thesis uses the consistent two-scales estimator of Zhang, Mykland and Aït-Sahalia (2005) with a slower frequency of ten ticks returns, which corrects for this source of bias.[7,8]

As mentioned previously, the motivation behind the HAR-RV model is the hypothesis that market participants are not homogeneous in behavior: they have different trading frequencies, and hence different reactions to the same news in the same market (leading to different volatility components). Corsi (2009) thus identifies three primary volatility components: short-term traders with trading frequency of a day or higher, medium-term investors that rebalance positions weekly, and participants with long-term horizons of one or more months. Realized volatility over time horizons longer than one day is calculated as the average of the daily measures across the time horizon, so that for a week and a month the measures are

[5] "Continue à droite, limite à gauche". That is, the process is right continuous with left limits.

[6] The quadratic variation and integrated variance coincide for the process in equation (1), but this no longer holds for more general return processes, as will be seen in Section 4.2.

[7] For a slower frequency S, realized variance can be calculated using different subsamples of ticks (e.g., for a frequency of 5 minutes, one can use prices at 9:30, 9:35, 9:40,... or 9:31, 9:36, 9:41,...). The two-scales measures combines the average of the subsample measures $RV_{t,1}, \ldots, RV_{t,S}$, together with the estimator $RV_t^{(all)}$ that uses all transaction returns within the day. Let n_t be the number of transactions on day t, the two-scales measure on that day is then given by $S^{-1}\sum_{i=1}^{S} RV_{t,i} - \frac{(n_t - S+1)/S}{n_t} RV_t^{(all)}$.

[8] The frequency is taken to be the same as in Corsi (2009) for replication purposes, but can also be chosen optimally (see, e.g., Russell and Bandi (2004) and Zhang et al. (2005)).

$$RV_t^{(w)} = \frac{1}{5}\sum_{i=0}^{4} RV_{t-i} \quad \text{and} \quad RV_t^{(m)} = \frac{1}{22}\sum_{i=0}^{21} RV_{t-i}, \tag{4}$$

respectively. The HAR-RV combines the terms in a simple time series representation

$$RV_{t+1} = c + \beta^{(d)} RV_t + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \omega_{t+1}, \tag{5}$$

which can be estimated with Ordinary Least Squares (OLS), yielding consistent and normally distributed estimates under standard assumptions.

To obtain h-day-ahead volatility forecasts for $h > 1$, the regressors need to be iteratively recalculated by including previously forecasted values. For instance, a 10-day-ahead forecast uses $RV_t^{(w)}$ and $RV_t^{(m)}$ based on 5 and 9 earlier predictions, respectively.

3.2 Empirical Analysis

3.2.1 Data

To derive the realized volatility, tick-by-tick price data of the Standard and Poors Depository Receipts (SPDR) S&P 500 Trust (ticker symbol SPY), the largest Exchange Traded Fund (ETF) on the S&P 500 index (VettaFi, 2024), is obtained from the NYSE Trade and Quote (TAQ) database. This asset is chosen due to its high liquidity: not only do investors use SPY to gain exposure to the overall direction of the market, but active traders frequently favor it as a means to switch between high-risk and low-risk assets (see, e.g., Elton, Gruber, Comer and Li (2005)). Regardless of liquidity, however, high-frequency data generally includes noisy observations (such as observations outside trading hours), which deteriorates the precision of realized volatility. Therefore, the high-frequency data cleaning procedure of Barndorff-Nielsen, Hansen, Lunde and Shephard (2009) is applied. The selected sample runs from January 2, 2013, up until December 30, 2022.

Figure 1 demonstrates the daily closing prices of the SPDR ETF in the sample period, and the corresponding daily two-scales realized volatility measure of Zhang et al. (2005). The ETF shows a clear uptrend over the entire sample period, with occasionally a strong downward move. There are two apparent spikes in volatility, which correspond to the COVID-19 pandemic at the beginning of 2020 and the Russian invasion of Ukraine in 2022. Finally, Table 1 provides summary statistics for the realized volatility across both the full sample and a subsample, which excludes the years 2020 and after. The realized volatility measure attains more extreme values in the full sample (as can be seen by the larger kurtosis). It can therefore be challenging to properly forecast volatility in the years 2020 and after, so that the forecasting exercise will only consider the sample excluding those years. However, a separate analysis of the full sample (including the years of turmoil) is done in Appendix A.

3.2.2 Forecasting Exercise

The predictive accuracy of the HAR-RV model is assessed by means of a horse race against benchmark models. The benchmarks considered are the AR(1), AR(3), and AR(22) for realized volatility, as well as a fractionally integrated model for realized volatility (Andersen et al., 2003).

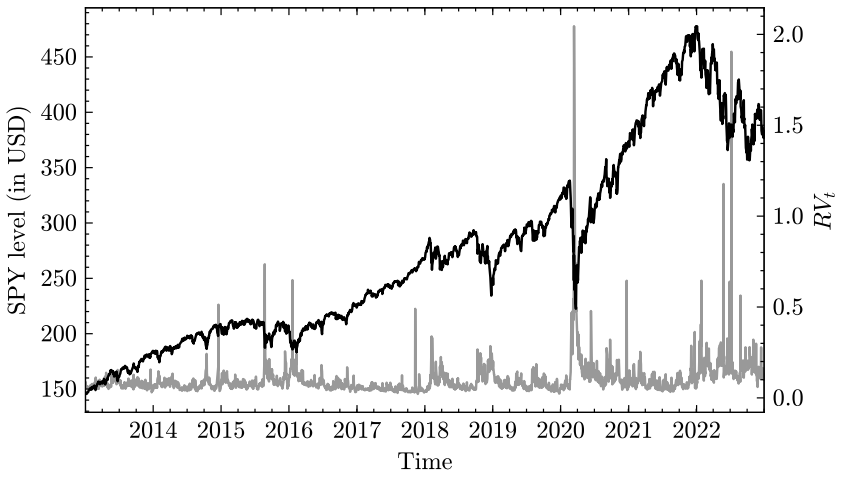


Figure 1: SPDR ETF (ticker symbol SPY) level running from January 2, 2013, up until December 30, 2022 (in black), and corresponding (annualized) two-scales realized volatility measure of Zhang et al. (2005) with a slower frequency of ten ticks returns (in gray).

Table 1: Summary Statistics of Realized Volatility.

	Mean	Std.dev.	Max.	Min.	Skew.	Kurt.
Sample period						
2013-2019	0.080	0.046	0.735	0.022	4.750	45.958
2013-2022	0.100	0.095	2.045	0.022	9.200	144.147

Notes: Summary statistics of the (annualized) realized volatility measure for a subsample (excluding the year 2020 and later) and the full sample (January 2, 2013, up until December 30, 2022), respectively. Realized volatility is calculated as proposed by Zhang et al. (2005), using a slower frequency of ten ticks returns.

The latter is estimated via a two-step procedure: first, the fractional coefficient d is estimated on the full sample using the GPH algorithm of Geweke and Porter-Hudak (1983), after which the fractional difference of the realized volatility series is taken, and fitted by an AR(5) model. Corsi (2009) shows that this ARFIMA$(5, d, 0)$ model is considered a good competitor of the HAR-RV model in terms of realized volatility forecasting performance.

Three forecasting horizons are considered: one day, one week, and two weeks. Forecasting performance for horizons longer than a day is evaluated based on the aggregated volatility over the horizon (i.e., the sum of realized volatility across the time period). To evaluate the forecasts, this thesis considers the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the R^2 of the Mincer-Zarnowitz (MZ) regressions of Mincer and Zarnowitz (1969). For h-day-ahead forecasts, the MZ regression becomes

$$\sum_{j=0}^{h} RV_{t+j} = \beta_0 + \beta_1 \mathbb{E}_{t-h}\left[\sum_{j=0}^{h} \widehat{RV}_{t+j}\right] + \varepsilon_t, \tag{6}$$

where $\mathbb{E}_{t-h}[\cdot]$ denotes the expectation given all information on day $t-h$ and $\widehat{RV}_{t+j}$ is the j-day-ahead forecast of realized volatility of one of the models. In the MZ regression, parameter values of $\beta_0 = 0$ and $\beta_1 = 1$ are a sign that efficient forecasts are produced (i.e., the forecast contains all information available at the time it is made), so that for completeness, the joint null

of $\beta_0 = 0$ and $\beta_1 = 1$ is tested for each MZ regression by using the Wald statistic (assuming the standard OLS assumptions hold so that estimates are asymptotically normal).

Table 2 shows the results for out-of-sample realized volatility forecasts. All models are re-estimated daily using a moving window of 1000 days (approximately 4 years) and the parameter d in the ARFIMA model is estimated on the full sample. Forecasts are then obtained using an AR(5) model fitted on the fractionally differenced series. For the dataset considered, a value of $d = 0.420$ is found, which is similar to Corsi (2009).

Looking at the RMSE and MAE, the HAR-RV model is found to outperform all benchmark models considered. Furthermore, the MZ regressions R^2 values for each forecast horizon are the highest for the HAR-RV model, highlighting the model's superiority in producing efficient forecasts. Table 3 further supports this, showing no rejection of the MZ joint null for all horizons (the AR(22) model also shows this). In contrast to Corsi (2009), the ARFIMA model turns out to be the worst performing model for each horizon considered. This is likely due to the difference in data (a sample from 1990 to 2007 was used), implying that the ARFIMA model is not well suited during the considered sample period.

As a final step, it is worth investigating whether the out-of-sample accuracy of the HAR-RV model is significantly higher than that of the benchmark models. This can be done by the Diebold-Mariano (DM) test of equal predictive accuracy (Diebold & Mariano, 2002). Table 4 shows the p-values corresponding to the modified DM statistics of Harvey, Leybourne and Newbold (1997), where the null hypothesis is that the HAR-RV model has the same MSE as the model considered in the corresponding row for the given forecasting horizon. At the 5% significance level, there are two points to mention. First, both the AR(1) and ARFIMA model are out of the competition, as the null is rejected for all horizons. Second, the AR(3) and AR(22) do not perform significantly worse than the HAR-RV model for horizons of one day and one week, but show a significantly higher MSE for the longest forecast horizon. Hence, it can be concluded that the HAR-RV model shows the best out-of-sample forecasting performance for longer horizons.[9] This is a useful property for the option pricing framework described in the following section. Further demonstration of favorable properties of the HAR-RV model are shown by means of a simulation study in Appendix B.

4 Option Pricing Methodology

This section describes the HAR option pricing framework proposed by this thesis. The steps of the framework are as follows: (i) A modeling assumption on the asset dynamics; (ii) A change in the probability measure; (iii) Discretization of risk-neutral dynamics for simulation. This thesis considers two different modeling assumptions in the first step (either with, or without jumps). For convenience, the terms "framework" and "model" are thus used interchangeably in the remainder of this thesis. After describing the models in detail, their assessment and benchmarks are discussed.

[9]For completeness, this thesis additionally performs the Model Confidence Set (MCS) procedure of Hansen, Lunde and Nason (2011). Just like the DM test, the MCS procedure investigates whether the models have equal predictive accuracy, but tends to perform slightly better in finite samples. Using a 5% significance level and a squared loss function, the HAR-RV model is in the MCS for all forecasting horizons. Inclusion of the AR(3) and AR(22) models across horizons resonates with the DM p-values in Table 4.

Table 2: Out-of-sample Forecasting Performance.

	$h=1$			$h=5$			$h=10$		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
AR(1)	0.033	0.020	0.538	0.163	0.121	0.508	0.338	0.265	0.402
AR(3)	0.032	0.019	0.574	0.138	0.095	0.568	0.289	0.213	0.478
AR(22)	0.032	0.019	0.569	0.137	0.089	0.566	0.274	0.190	0.488
ARFIMA	0.090	0.079	0.429	0.440	0.398	0.415	0.876	0.797	0.350
HAR	0.031	0.018	0.576	0.135	0.089	0.571	0.271	0.189	0.497

Notes: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the R^2 of the MZ regressions of the AR(1), AR(3), ARFIMA$(5, d, 0)$, and HAR-RV models for the SPDR ETF over the sample January 2, 2013, up until December 30, 2019. Forecasting horizons of one day, one week, and two weeks are considered ($h = 1$, 5, and 10, respectively). The AR(1), AR(3), and HAR-RV models are re-estimated daily using a moving window of 1000 observations. This is also the case for the ARFIMA model, the only difference being that prior to the forecasting exercise, the parameter d is estimated on the full sample and used to take the fractional difference of the realized volatility series. Forecasts are then obtained using an AR(5) model on the fractionally differenced series.

Table 3: MZ regression Wald test p-values.

	$h=1$	$h=5$	$h=10$
AR(1)	0.000	0.000	0.000
AR(3)	0.011	0.000	0.000
AR(22)	0.359	0.144	0.294
ARFIMA	0.000	0.000	0.000
HAR	0.134	0.105	0.271

Notes: p-values of the Wald test statistics for the joint null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ in the MZ regressions of Table 2.

Table 4: DM test p-values.

	HAR		
	$h=1$	$h=5$	$h=10$
AR(1)	0.048	0.000	0.000
AR(3)	0.478	0.076	0.000
AR(22)	0.227	0.173	0.007
ARFIMA	0.000	0.000	0.000

Notes: DM test p-values for the null of equal forecast accuracy (accuracies in Table 2) between the HAR-RV model and the model in the corresponding row, for a given forecasting horizon. The alternative is one sided (that is, the MSE of the HAR-RV model is smaller than the MSE of the other competing model. The p-values correspond to the modified DM statistics of Harvey et al. (1997).

4.1 HAR-RV Option Pricing Model

For the first model, the dynamics of the logarithmic asset price are assumed to satisfy the same SDE as in equation (1). In addition, the asset is assumed to continuously pay out a constant dividend yield over the lifetime of the option.

4.1.1 Change of Probability Measure

To price financial products, it should hold that they are free of arbitrage, that is, no risk-free profits should be present. Prior to pricing, the following theorem should be introduced (which is also required for the second option pricing model). Let the probability measure in which equation (1) holds be denoted by $\mathbb{P}$ (the so-called "physical measure"). The Fundamental Theorem of Asset Pricing states that a market on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is free of arbitrage if, and only if, there exists a probability measure $\mathbb{Q}$ that is equivalent[10] to $\mathbb{P}$, and any discounted value process $V(t)$ is a martingale under $\mathbb{Q}$, that is,

[10]Two probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ with sample space Ω are equivalent if, and only if, $\mathbb{P}(A) > 0$ when $\tilde{\mathbb{P}}(A) > 0$ and also the other way around, so $\tilde{\mathbb{P}}(A) > 0$ when $\mathbb{P}(A) > 0$, for all $A \subseteq \Omega$.

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}V(T)\middle|\mathcal{F}(t)\right] = V(t), \quad 0 \leq t \leq T, \tag{7}$$

where $\mathbb{E}^{\mathbb{Q}}[\cdot]$ denotes the expectation under $\mathbb{Q}$, the so-called "risk-neutral measure". In other words, the current value of an asset should be its discounted expected value (referred to as the martingale property[11]) in order for it to be arbitrage-free. This is not the case under the physical probability measure, and hence a change of measure (from $\mathbb{P}$ to $\mathbb{Q}$) needs to be performed. Using Girsanov's theorem, one can show that $X(t)$ is a martingale under $\mathbb{Q}$ when it has the dynamics

$$dS(t) = (r-q)S(t)dt + \sigma(t)S(t)dW^{\mathbb{Q}}(t), \tag{8}$$

where r is the risk-free interest rate, q is the dividend yield, and $W^{\mathbb{Q}}(t)$ is the Brownian motion under the risk-neutral measure.

4.1.2 Monte Carlo Procedure

At a day t_0, the Monte Carlo procedure to price an option with time to maturity $\tau = T - t_0$, strike price K, and underlying with risk-neutral dynamics as in equation (8), is as follows. First, the time interval until expiration $[t_0, T]$ is partitioned into days $t_0 < t_1 < \cdots < T$. Denote the time interval between two days by $\Delta t = t_{i+1} - t_i$. As a next step, let P be the number of asset paths that will be generated and let $s_{i,j} = S(t_i)$ for $j = 1, \ldots, P$. To simulate the asset paths, the continuous-time process in equation (8) must be discretized. As is common in literature, this thesis applies an Euler scheme to do so (Butcher, 2016), which yields a discretization of the form

$$s_{i+1,j} \approx s_{i,j} + (r-q)s_{i,j}\Delta t + \sigma(t_i)s_{i,j}W(\Delta t), \tag{9}$$

where $\sigma(t_i)$, the volatility on day t_i, is approximated by realized volatility predictions of the HAR-RV model. That is, on day t_0, predictions are made for days $t_1, \ldots, T$ as in Section 3, using the past 1000 observations to estimate the parameters $\beta^{(d)}$, $\beta^{(w)}$ and $\beta^{(m)}$.[12] Furthermore, note that by the properties of the Brownian motion, it holds that $W(\Delta t) \sim N(0, \Delta t)$. Hence, $W(\Delta t)$ in equation (9) can be replaced by $\sqrt{\Delta t}Z$ with $Z \sim N(0,1)$, which can easily be simulated.

After obtaining all asset paths, the payoff of each path can be calculated. In particular, let H_j be the option payoff of path $j = 1, \ldots, P$. As this thesis only considers European options, the payoffs are given by

$$H_j^c = \max(s_{T,j} - K, 0) \quad \text{and} \quad H_j^p = \max(K - s_{T,j}, 0), \tag{10}$$

for call and put options, respectively. The final step is deriving the approximated option value

[11] A value process, or any stochastic process $Y(t)$ that is adapted (i.e., $Y(t)$ is $\mathcal{F}(t)$-measurable) and integrable (i.e., $\mathbb{E}^{\mathbb{P}}[|Y(t)|] < \infty$, $\forall\, t \geq 0$) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a martingale if $\mathbb{E}^{\mathbb{P}}[Y(t)|\mathcal{F}(s)] = Y(s)$, $\forall\, s \leq t$.

[12] The options considered in this thesis are so-called "PM-settled": they expire after the market close on the expiration date. Hence, a volatility forecast for the full trading session on the expiration day is required.

at time t_0, which is the discounted average payoff over all paths, that is,

$$V(t_0) \approx e^{-r\tau} \frac{1}{P} \sum_{j=1}^{P} H_j. \tag{11}$$

Theoretically, the previously described procedure should yield an option price close to the market price. However, an issue might arise in the forecasting of realized volatility with the HAR-RV model: as it is equivalent to a restricted AR model, predictions will converge to the conditional mean when the forecasting horizon increases. Therefore, the following subsection introduces a HAR option pricing model which addresses this pseudo-constant volatility over the options' remaining lifetimes by including random jumps.

4.2 HAR-BPV Option Pricing Model

Alternatively, one can consider a jump-diffusion process for $X(t)$, that is,

$$dX(t) = \mu(t)dt + \sigma(t)dW(t) + J(t)dX_{\mathcal{P}}(t), \quad 0 \leq t \leq T, \tag{12}$$

where $X_{\mathcal{P}}(t)$ is a Poisson process with intensity parameter $\xi > 0$ and J is the jump magnitude, which follows a distribution F_J (e.g., Gaussian, Laplace, ...). Furthermore, $W(t)$ and $X_{\mathcal{P}}(t)$ are assumed to be independent. The daily quadratic variation of the return process is different than in equation (2), as it now includes the jump component, that is,

$$QV_t = \int_{t-1}^{t} \sigma^2(s)\, ds + \sum_{t-1 \leq s \leq t} J^2(s), \tag{13}$$

with the first term being the integrated variance and the second the cumulative squared jump magnitudes. In this setting, RV_t^2 still converges in probability to QV_t for increasing sampling frequency and under absence of microstructure noise (Barndorff-Nielsen & Shephard, 2002).

As in the previous subsection, the goal is to discretize the dynamics under the risk-neutral measure and perform Monte Carlo simulations. The question then arises on how to separate the two terms in equation (13). Barndorff-Nielsen and Shephard (2004) defines the so-called "bipower variation", a jump-robust measure for the integrated variance, as

$$BPV_t = \sqrt{\mu^{-2} \sum_{i=2}^{M} |r_{t_i}||r_{t_{i-1}}|} \tag{14}$$

where $\mu^{-2} = \sqrt{2/\pi}$ is the absolute first moment of a standard normally distributed random variable. Furthermore, they show that for increasing sampling frequency, this term converges in probability to the integrated variance, that is,

$$BPV_t^2 \xrightarrow{p} \int_{t-1}^{t} \sigma^2(s)\, ds, \tag{15}$$

and hence, one can obtain the daily jump component as

$$RV_t^2 - BPV_t^2 \xrightarrow{p} \sum_{t-1 \leq s \leq t} J^2(s), \tag{16}$$

which is truncated to be positive as suggested by Barndorff-Nielsen and Shephard (2004). Therefore, after calculating the daily realized variance and bipower variation (as before, using the returns sampling frequency of ten ticks), an estimate of the daily jump component is obtained by taking the square root of their difference. The sign of the jump is then derived from the sign of the return on the corresponding day.

4.2.1 Change of Probability Measure

Again, by the Fundamental Theorem of Asset Pricing, the underlying dynamics must be a martingale under $\mathbb{Q}$ in order to be used to price derivatives on it. It can be proven (see Appendix C) that under the risk-neutral measure, the dynamics in equation (12) become

$$\frac{dS(t)}{S(t)} = \left(r - q - \xi \mathbb{E}\left[e^J - 1\right]\right) dt + \sigma(t) dW^{\mathbb{Q}}(t) + (e^J - 1) dX_{\mathcal{P}}^{\mathbb{Q}}(t), \tag{17}$$

where $W^{\mathbb{Q}}(t)$ and $X_{\mathcal{P}}^{\mathbb{Q}}(t)$ are the Brownian motion and Poisson process under the risk-neutral measure, respectively. Compared to equation (8), the drift term now contains a so-called "drift compensation term" equal to $\xi \mathbb{E}\left[e^J - 1\right]$, which corrects for the jump component in order to make the process a martingale under $\mathbb{Q}$.

4.2.2 Monte Carlo Procedure

The Monte Carlo option pricing procedure is analogous to the one where no jump-diffusion process is assumed. There are some slight differences, such as the discretization. Again, an Euler scheme is used, which now results in the form

$$s_{i+1,j} \approx s_{i,j} + (r - q - \xi \mathbb{E}\left[e^J - 1\right]) s_{i,j} \Delta t + \sigma(t_i) s_{i,j} W(\Delta t) + (e^J - 1) s_{i,j} X_{\mathcal{P}}(\Delta t), \tag{18}$$

for $t_i = t_1, \ldots, T$ and $j = 1, \ldots, P$. As the volatility $\sigma(t_i)$ is now assumed to not contain any jumps, it should be approximated by the bipower variation, rather than by the realized volatility. The HAR-RV model is adjusted to the HAR-BPV model, that is,

$$BPV_{t+1} = c + \beta^{(d)} BPV_t + \beta^{(w)} BPV_t^{(w)} + \beta^{(m)} BPV_t^{(m)} + \omega_{t+1d}, \tag{19}$$

where the weekly and monthly terms are calculated in similar fashion to the monthly and weekly measures for realized volatility. Estimation and forecasting is analogous to the HAR-RV model.

It now remains to approximate the jump component in equation (18), which requires knowledge on the Poisson parameter ξ and the jump distribution F_J. A simple idea to estimate ξ is to count the number of jumps that occur over a time period. However, it could occur that the difference between the realized variance and the bipower variation is non-zero due to sampling variation, rather than due to a jump. Hence, one should only consider "significant jumps", which can be identified with the test introduced by Barndorff-Nielsen and Shephard (2004) at

a chosen significance level (Appendix D provides a detailed explanation of the test). Counting the significant jumps over a time period yields an estimate $\hat{\xi}$, which, together with the property that $X_{\mathcal{P}}(\Delta t)$ is Poisson distributed with parameter $\xi \Delta t$, can be used to draw random variables.

Using the significant jumps, an assumption can be made on the distribution of their magnitudes. Jumps are often found to be asymmetric in size: negative jumps are generally larger than positive jump. Hence, this thesis assumes a nonsymmetric double exponential distribution for the jump magnitude, also known as Kou's model (Kou, 2002). The density is defined as

$$f_J(x) = \begin{cases} p_1 \cdot \alpha_1 e^{-\alpha_1 x}, & \text{if } x \geq 0 \\ p_2 \cdot \alpha_2 e^{\alpha_2 x}, & \text{otherwise,} \end{cases} \tag{20}$$

where $p_1, p_2 \geq 0$ and $p_1 + p_2 = 1$.[13] Using the daily jumps obtained as explained below equation (16), parameter estimates are found with MLE.

Summarizing, the volatility $\sigma(t_i)$ is obtained from the HAR-BPV forecasts, the intensity parameter ξ is based on counting the significant jumps, and the parameters of f_J are estimated with MLE on the magnitudes of these jumps. It now remains to calculate the payoffs of the simulated asset paths at maturity, from which the option value estimate on day t_0 can be derived.

4.3 Assessing Model Performance

4.3.1 Benchmark Models

This thesis considers two benchmarks, which are either based on the BS model, or the GARCH option pricing model of Heston and Nandi (2000). The BS model is "plug-and-play", that is, a formula with inputs is provided from which the call option price can easily be derived (the value for put options is obtained by using the put-call parity). Although strong from a theoretical perspective, the BS model fails to exhibit the volatility smile. Hence, this thesis considers the Practitioners BS (PBS) model proposed by Dumas et al. (1998), where each option has a volatility which is a function of the strike price and time to maturity. The considered functional form is

$$\sigma_{\text{BS},i} = \beta_0 + \beta_1 K_i + \beta_2 K_i^2 + \beta_3 \tau_i + \beta_4 \tau_i^2 + \beta_5 K_i \tau_i + \nu_i, \tag{21}$$

where $\sigma_{\text{BS},i}$ is the BS implied volatility in option i. Equation (21) is estimated on each observation day with OLS, using all options observed in the past week. Although not theoretically consistent, the PBS model is more realistic than the "as is" BS model, since it uses information from other options. In addition, Dumas et al. (1998) show that it outperforms several popular option pricing models, such as the deterministic volatility model of Dupire et al. (1994). It is therefore expected to be a challenging benchmark for the proposed HAR option pricing models.

As this thesis proposes a model where time is discretized, a natural competitor is the GARCH(1,1) option pricing model of Heston and Nandi (2000). They present an asymmetric GARCH model, where $X(t)$ is assumed to follow a discrete-time process under the physical measure $\mathbb{P}$ that is defined as

[13] To derive the expectation in equation (18), it should additionally hold that $\alpha_1 > 1$ and $\alpha_2 > 0$ in order to integrate e^x, which yields $\mathbb{E}[e^J] = p_1 \cdot \frac{\alpha_1}{\alpha_1 - 1} + p_2 \cdot \frac{\alpha_2}{\alpha_2 + 1}$.

$$
\begin{aligned}
X(t) &= X(t-1) + r + \lambda h(t) + \sqrt{h(t)} z(t), \\
h(t) &= \omega + \beta h(t-1) + \alpha \cdot \left(z(t-1) - \gamma \sqrt{h(t-1)} \right)^2,
\end{aligned}
\tag{22}
$$

where $h(t)$ is the conditional variance at time t and $z(t) \sim N(0,1)$. The conditional variance in equation (22) is different from the standard GARCH of Bollerslev (1986) due to the γ-term, making it more related to the nonlinear GARCH model of Engle and Ng (1993). The parameter γ captures the leverage effect (the negative correlation between returns and volatility), whereas the parameter λ captures the equity risk premium. Under $\mathbb{Q}$, the form in equation (22) remains the same, with the exception of the parameters λ and γ being replaced by $\lambda^* = -1/2$ and $\gamma^* = \gamma + \lambda + 1/2$. Parameters are estimated with MLE, after which options can be priced. Heston and Nandi (2000) show that the closed-form for the call option price is given by

$$
\begin{aligned}
V(t_0) = \frac{1}{2} S(t) + \frac{e^{-r\tau}}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{K^{-i\phi} f^*(i\phi + 1)}{i\phi} \right] d\phi \\
- K e^{-r\tau} \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{K^{-i\phi} f^*(i\phi)}{i\phi} \right] d\phi \right),
\end{aligned}
\tag{23}
$$

where $\mathrm{Re}[\cdot]$ is the real part of a complex number and $f(\phi) = \mathbb{E}^{\mathbb{P}}[e^{X(T)\phi} | \mathcal{F}(t)] = S(t)^{\phi} e^{A(t) + B(t) h(t+1)}$ is the conditional moment generating function of the logarithmic asset price (and $f^*(\phi)$ is the version under $\mathbb{Q}$). The coefficients $A(t)$ and $B(t)$ can be calculated recursively from the terminal condition $A(T) = B(T) = 0$ by using that

$$
\begin{aligned}
A(t) &= A(t+1) + \phi r + B(t+1)\omega - \frac{1}{2} \log(1 - 2\alpha B(t+1)), \\
B(t) &= \phi(\lambda + \gamma) - \frac{1}{2}\gamma^2 + \beta B(t+1) + \frac{(\phi - \gamma)^2 / 2}{1 - 2\alpha B(t+1)}.
\end{aligned}
\tag{24}
$$

4.3.2 Performance Evaluation

The assessment of the option pricing models and their benchmarks is done as is customary in literature. That is, after valuating a number of options, the Root Mean Square Error on the percentage option implied volatility (RMSE_{IV}) is calculated as

$$
\mathrm{RMSE}_{IV} = \left(\frac{1}{N} \sum_{i=1}^{N} \left(IV_i^{\mathrm{mkt}} - IV_i^{\mathrm{mod}} \right)^2 \right)^{\frac{1}{2}},
\tag{25}
$$

where N is the number of options, and IV^{mkt} and IV^{mod} are the BS implied volatilities for the market and model option prices, respectively. One could also consider the price RMSE (RMSE_P), where the implied volatilities are replaced by the option prices. However, Christoffersen and Jacobs (2004) show that the RMSE_P is particularly relevant for at-the-money (ATM) options (as these are more expensive), whereas the RMSE_{IV} gives more weight to OTM options. Since this thesis only considers OTM options (for reasons described in Section 5), the RMSE_{IV} is preferred.

5 Empirical Analysis

In this section, the HAR option pricing models described in the previous section are used to price weekly and monthly option on the S&P 500 index. Furthermore, the models' performances are assessed by comparing to the PBS and GARCH benchmarks.

5.1 Data

The option data is obtained from OptionMetrics[14] and consists of S&P 500 index[15] options written between January 2, 2013, and December 30, 2022 (the same sample as in Section 3). Within maturities (τ, in days), three categories are considered: "weekly" ($\tau \leq 9$), "biweekly" ($9 < \tau \leq 20$), "monthly" ($20 < \tau \leq 60$), and "long-term" ($60 < \tau \leq 180$). Although the latter category is not the aim of this thesis, it is interesting for comparison with literature. As is customary, options are observed on Wednesdays to avoid seasonal effects from Mondays and Fridays as much as possible. Furthermore, only OTM options are considered as these are the most liquid and have tighter spreads, which results in more accurate option prices.[16] In addition, this thesis partially implements the options data cleaning procedure of Andersen et al. (2017) and several other filters for low liquidity contracts. Appendix E describes these.

For a structured analysis, options within each maturity group are split up into different "moneyness" categories, which represents the distance of the underlying asset price to the strike price. As in Andersen et al. (2017), this thesis uses the so-called "standardized (forward) moneyness", which, for an option with strike price K and τ days to maturity, is defined as

$$m = \frac{\ln(K/F_\tau)}{\sqrt{\tau}\, IV_{\text{ATM},\tau}}, \tag{26}$$

where $F_\tau = S(t)e^{r\tau}$ is the forward price of the underlying for τ days and $IV_{\text{ATM},\tau}$ is the implied volatility of the option with strike price nearest to F_τ. The standardized moneyness can be interpreted as the number of standard deviations the forward price is above the strike price. As this thesis omits ITM calls and puts, it holds that $m > 0$ and $m < 0$ correspond to OTM calls and OTM puts, respectively. Five categories of moneyness are defined: deep-OTM puts ($m \leq -3$), OTM puts ($-3 < m \leq -1$), ATM options ($-1 < m \leq 1$), OTM calls ($1 < m \leq 3$), and deep-OTM calls ($3 < m$). In the following, options that are not deep-OTM are referred to as non-extreme moneyness options.

After cleaning, the dataset contains a total of 2,744,594 options. Distributing them across categories yields the structure in Table 5, where the average implied volatility, standard deviations of implied volatility, and the number of contracts for each category are reported. It is clear that the implied volatility is not constant across moneyness and maturity, showing the volatility smile for longer maturities. This implies that the underlying asset return distribution is not

[14] OptionMetrics additionally provides the risk-free rate (Zero-Coupon yield curve) and the dividend yield of the index, which are both used in this thesis.

[15] Note that the S&P 500 index is not equivalent to the SPDR ETF, from which the realized volatility is calculated. However, the daily index and ETF level series have a correlation of 1 in the sample, so that this thesis assumes equivalence between the two. It can easily be shown that their realized volatilities are the same under this assumption.

[16] In-the-money (ITM) options have a large "delta", which is the sensitivity of the option price with respect to the underlying asset price. Such positions are more costly to hedge, making them less appealing for traders.

symmetric and exhibits fat tails, and confirms the use of time-varying volatility models. Furthermore, the implied volatility shows the most variation for OTM puts with short maturities, which decreases for higher moneyness levels and longer maturities.

Due to the large sample size, it is computationally too expensive to price all of the options. Hence, this thesis focuses on two subsamples (that also ensure enough training data for the models is available): the year 2018 (320,597 observations) and the year 2022 (517,110 observations). Section 3 showed that these two periods have totally different volatility profiles, making this an interesting robustness check for the proposed models. For the analyses in 2018 and 2022, models are trained using a subsample (2013-2018) and full sample (2013-2022), respectively.

Table 5: Options Dataset.

	Maturity			
	$\tau \leq 9$	$9 < \tau \leq 20$	$20 < \tau \leq 60$	$60 < \tau \leq 180$
Implied volatility				
$m \leq -3$	0.399	0.386	0.389	0.416
$-3 < m \leq -1$	0.287	0.254	0.257	0.284
$-1 < m \leq 1$	0.224	0.198	0.188	0.197
$1 < m \leq 3$	0.190	0.157	0.149	0.151
$3 < m$	0.176	0.167	0.179	0.208
Implied volatility std. dev.				
$m \leq -3$	0.194	0.164	0.157	0.155
$-3 < m \leq -1$	0.170	0.114	0.098	0.081
$-1 < m \leq 1$	0.154	0.109	0.095	0.072
$1 < m \leq 3$	0.105	0.081	0.069	0.049
$3 < m$	0.074	0.074	0.070	0.049
Number of contracts (%)				
$m \leq -3$	2.850	3.923	12.768	10.539
$-3 < m \leq -1$	1.784	2.065	8.482	11.226
$-1 < m \leq 1$	1.996	2.649	11.146	16.847
$1 < m \leq 3$	1.500	1.649	5.387	4.451
$3 < m$	0.120	0.193	0.289	0.137

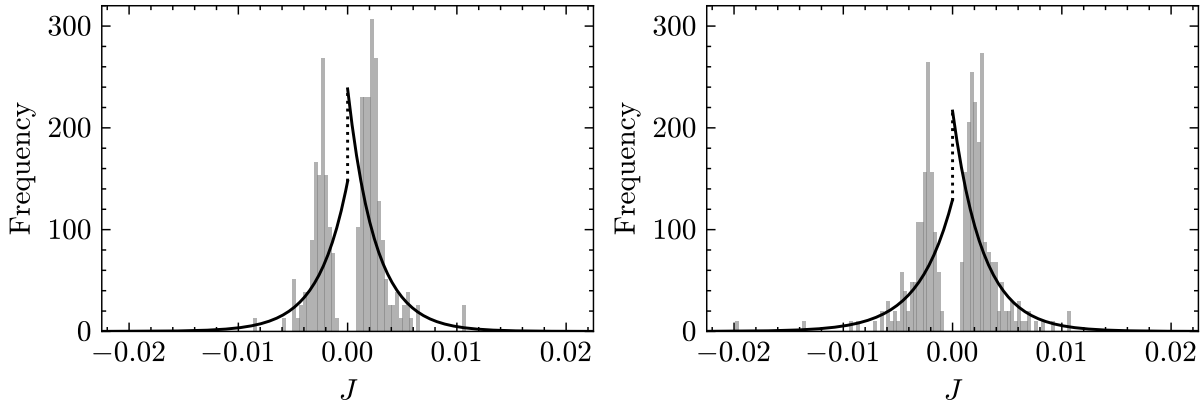
Notes: Average implied volatility (upper), standard deviation of implied volatility (middle), and number of option contracts in each moneyness-maturity category (lower) for the option dataset. The options are out-of-the-money and are written on the S&P 500 index. The sample runs from January 2, 2013, up until December 30, 2022, where options are only observed on Wednesdays, yielding a total of 2,744,594 observations. Implied volatility is calculated with the Black-Scholes model. Moneyness (m) is the standardized (forward) moneyness and maturity (τ) is in calendar days.

5.2 Option Pricing Results

Prior to analyzing the performance of the option pricing models in the years 2018 and 2022, some preparatory results are discussed. Figure 2 shows the estimated nonsymmetric double exponential distribution for both the subsample and full sample, where parameters have been estimated with MLE using only jumps that are significant at a level of 1%. The gap around zero is due to the exclusion of nonsignificant jumps. Clearly, negative and positive jumps have

different magnitudes, and hence a symmetric distribution would not be appropriate.[17] Dividing the number of significant jumps by the number of trading days yields estimates for the Poisson parameter ξ, which is found to be 0.137 and 0.133 in the subsample and full sample, respectively.

Finally, Table 6 shows the parameter estimates of both the jump magnitude distribution and the GARCH model of Heston and Nandi (2000) for the sample up until 2018, and the full sample. Surprisingly, the estimates of the jump distribution parameters do not differ much between the two samples. The largest change is in the left side of the density (represented by α_2) due to the large negative jumps that arise in the years of turmoil. The GARCH model estimates show greater contrasts: the leverage effect parameter γ is larger for the full sample, implying that the correlation between negative returns and volatility is more negative when including the years of turmoil. In addition, the equity risk premium parameter λ is three times as large in the full sample, reflecting the higher reward an investor should get for taking more risk.



(a) Reduced Sample

(b) Full Sample

Figure 2: Illustration of jump magnitude distribution in the sample up until 2018 (2a), and the full sample (2b). The gray histogram illustrates the empirical distribution of the (significant) jumps, whereas the black line represents the estimated nonsymmetric double exponential distribution. Parameter estimates are from Table 6.

Table 6: Maximum Likelihood Estimates.

HAR-BPV	f_J parameters		
	p_1	α_1	α_2
2013-2018	0.606	393.299	374.364
2013-2022	0.594	364.056	319.512

GARCH					
	ω	α	β	γ	λ
2013-2018	4.155e-7	9.148e-6	0.792	90.029	0.495
2013-2022	3.895e-8	8.596e-6	0.752	139.591	1.537

Notes: MLE results for the HAR-BPV jump magnitude distribution and the GARCH model of Heston and Nandi (2000) for the subsample (2013-2018) and full sample (2013-2022). A constant risk-free interest rate of 1% per year has been used to obtain the GARCH parameters.

[17]Indeed, when considering a normal jump magnitude distribution in the spirit of Merton (1976), a much smaller log-likelihood function value is obtained (values of 470.990 and 1388.344 over the full sample for normal and nonsymmetric double exponential, respectively).

5.2.1 Calm Sample Period

The RMSE_{IV} of the HAR option pricing models and their benchmark for the year 2018 are shown in Table 7. For the HAR-RV option pricing model, each option has been priced using 10,000 simulated asset paths, while the HAR-BPV version uses ten times this amount, i.e. 100,000 simulations. The reason for this is that the jump intensity is relatively small, so that a large amount of paths are needed to obtain a reasonable amount of jump occurrences.[18]

The performance of the HAR option pricing models in Panels A and B show that adding a random jump component contributes to precision, but not in all cases. In particular, the HAR-RV option pricing model performs better than the HAR-BPV option pricing model only for deep-OTM calls and puts. An explanation for this could be that such options are tautologically cheaper than options closer to the ATM level, so that jump occurrences can quickly overprice options. Another point of attention is that, contrary to the standard deviation pattern in Table 5, the HAR option pricing models generally obtain the highest accuracy for biweekly options, rather than for long-term maturities. A similar result is found in Corsi et al. (2013), where their HAR option pricing model yielded the best performance in this maturity spectrum.[19]

Comparing the HAR option pricing models with the PBS benchmark in Panel C confirms that the assumption of a constant volatility over the option lifetime is not realistic: with the exception of some moneyness levels for longer maturities the HAR option pricing models consistently outperform the PBS model. In particular, the HAR-RV (HAR-BPV) option pricing model yields average improvements of around 20% (9%) and 10% (16%) for deep-OTM options and non-extreme moneyness options, respectively. The better performance of the PBS model for long-term options is likely due to deterioration of forecasting accuracy of the HAR model for longer horizons, as was shown in Section 3.

Panel D shows that the GARCH option pricing model of Heston and Nandi (2000) is a more competitive benchmark than the PBS model in Panel C. A minor improvement over the GARCH model by the PBS model is obtained for long-term options. This is consistent with Heston and Nandi (2000), where the PBS model was also found to have a higher accuracy for longer maturities at some moneyness levels.

Looking at Panels A,B, and D, both HAR option pricing models fail to outperform the GARCH model by a large amount for deep-OTM put options, implying that the latter is better in capturing the left tail of the future volatility. This is likely because the GARCH model incorporates the leverage effect (through the parameter γ), whereas the HAR models do not. However, both HAR option pricing models are highly competitive for near-ATM and OTM calls over all maturities, with the HAR-BPV option pricing model showing an average improvement of around 6% over the GARCH model in these categories. Corsi et al. (2013) also found the highest improvement for these moneyness levels. Finally, when only considering weekly and biweekly options, it can be concluded that the GARCH model should be used for deep-OTM options, whereas the HAR models are better at pricing non-extreme moneyness options.

[18] Oosterlee and Grzelak (2019) suggest simulating ten times more paths when rarely occurring jumps are added to a stochastic process.

[19] As mentioned earlier, Corsi et al. (2013) do not consider options with $\tau \leq 9$, so that this result is not entirely comparable.

Table 7: RMSE_{IV} of Option Pricing Models in 2018.

	Maturity			
	$\tau \leq 9$	$9 < \tau \leq 20$	$20 < \tau \leq 60$	$60 < \tau \leq 180$
Panel A: HAR-RV				
$m \leq -3$	0.120	0.092	0.109	0.118
$-3 < m \leq -1$	0.126	0.099	0.101	0.117
$-1 < m \leq 1$	0.065	0.047	0.044	0.048
$1 < m \leq 3$	0.045	0.028	0.025	0.018
$3 < m$	0.030	0.021	0.022	0.024
Panel B: HAR-BPV				
$m \leq -3$	0.129	0.100	0.111	0.119
$-3 < m \leq -1$	0.123	0.094	0.101	0.120
$-1 < m \leq 1$	**0.050**	**0.038**	**0.040**	0.050
$1 < m \leq 3$	**0.040**	**0.024**	**0.020**	**0.015**
$3 < m$	0.044	0.026	0.026	0.030
Panel C: PBS				
$m \leq -3$	0.197	0.145	0.134	**0.089**
$-3 < m \leq -1$	0.133	0.096	0.079	**0.055**
$-1 < m \leq 1$	0.092	0.073	0.063	0.048
$1 < m \leq 3$	0.077	0.052	0.042	0.028
$3 < m$	0.075	0.042	0.031	0.017
Panel D: GARCH				
$m \leq -3$	**0.066**	**0.076**	**0.084**	0.108
$-3 < m \leq -1$	**0.069**	**0.075**	**0.078**	0.099
$-1 < m \leq 1$	0.070	0.056	0.049	**0.036**
$1 < m \leq 3$	0.054	0.045	0.040	0.041
$3 < m$	**0.026**	**0.020**	**0.018**	**0.015**

Notes: Option pricing performance of the proposed option pricing models and their benchmarks on S&P 500 index options in the year 2018, observed on all Wednesdays (320,597 observations). The performance measure is the percentage implied volatility Root Mean Squared Error (RMSE_{IV}), where Black-Scholes implied volatilities are used. Measures are sorted by standardized moneyness (m) and calendar days to expiration (τ). Values in bold indicate the lowest RMSE_{IV} for the corresponding category of moneyness and maturity among all models. Parameter estimates for the jump magnitude in the HAR-BPV model and the GARCH model are as in Table 6.

5.2.2 Robustness Check: High Volatility Period

Table 8 shows the RMSE_{IV} for the HAR option pricing models and their benchmarks in the year 2022. The same amount of simulations has been used as in the year 2018, and parameter estimates for the jump dynamics and GARCH model are as in Table 6.

At first glance, there is a clear deterioration in performance compared to 2018 for all models across categories, implying that options are harder to price in periods of higher volatility. Interestingly, the HAR-RV option pricing model in Panel A now consistently beats its jump equivalent in Panel B, which was not the case in the year 2018. This seems rather surprising, as one would expect jumps to be required to mimic the high volatility in this period. However, as stated earlier, the estimated jump arrival rate ξ and jump distribution parameters in the full sample do not differ much from the sample without the year of turmoil, so that the jump behavior did not change much. Hence the lower performance of jump model can only be explained

by the deterioration in the predictive performance of the bipower variation, which seems to be larger than that of the realized volatility.

Making a comparison with the benchmarks shows that the HAR option pricing models are not competitive for pricing put options, as the GARCH and PBS models in Panels C and D attain the highest accuracy for deep-OTM puts and OTM puts, respectively. For ATM and call options, however, the HAR models yield lower RMSE_{IV} than both benchmarks, where the HAR-RV option pricing model even attains higher accuracy than in the year 2018. This shows that the models proposed by this thesis capture the right tail of the volatility distribution relatively well in periods of economic turmoil. Finally, the weekly and biweekly options are now best priced with the HAR-RV option pricing model if they are ATM or call options, whereas put options are best priced with the GARCH benchmark.

Table 8: RMSE_{IV} of Option Pricing Models in 2022.

	Maturity			
	$\tau \leq 9$	$9 < \tau \leq 20$	$20 < \tau \leq 60$	$60 < \tau \leq 180$
Panel A: HAR-RV				
$m \leq -3$	0.143	0.122	0.153	0.196
$-3 < m \leq -1$	0.138	0.131	0.160	0.176
$-1 < m \leq 1$	**0.093**	**0.075**	0.084	0.088
$1 < m \leq 3$	**0.068**	**0.042**	0.039	0.030
$1 < m \leq 3$	**0.030**	**0.025**	0.021	**0.015**
Panel B: HAR-BPV				
$m \leq -3$	0.148	0.126	0.172	0.196
$-3 < m \leq -1$	0.149	0.141	0.166	0.188
$-1 < m \leq 1$	0.099	0.083	0.082	0.114
$1 < m \leq 3$	0.074	0.047	**0.033**	**0.018**
$3 < m$	0.044	0.034	0.032	0.028
Panel C: PBS				
$m \leq -3$	0.182	0.167	0.226	0.214
$-3 < m \leq -1$	0.100	**0.072**	**0.079**	**0.060**
$-1 < m \leq 1$	0.097	0.075	**0.066**	**0.054**
$1 < m \leq 3$	0.084	0.064	0.053	0.032
$3 < m$	0.099	0.108	0.067	0.046
Panel D: GARCH				
$m \leq -3$	**0.103**	**0.066**	**0.068**	**0.084**
$-3 < m \leq -1$	**0.085**	0.104	0.103	0.110
$-1 < m \leq 1$	0.105	0.084	0.078	0.092
$1 < m \leq 3$	0.074	0.057	0.040	0.031
$3 < m$	0.048	0.035	**0.019**	0.021

Notes: Option pricing performance of the proposed option pricing models and their benchmarks on S&P 500 index options in the year 2020, observed on all Wednesdays (517,110 observations). The performance measure is the percentage implied volatility Root Mean Squared Error (RMSE_{IV}), where Black-Scholes implied volatilities are used. Measures are sorted by standardized moneyness (m) and calendar days to expiration (τ). Values in bold indicate the lowest RMSE_{IV} for the corresponding category of moneyness and maturity among all models. Parameter estimates for the jump magnitude in the HAR-BPV model and the GARCH model are as in Table 6.

6 Conclusion

This thesis presents a simple option pricing framework based on the HAR model of Corsi (2009). First, the capability of the HAR-RV model to predict realized volatility of the S&P 500 SPDR ETF was assessed by means of a horse race against AR and ARFIMA benchmarks. The superiority of the HAR-RV model shown by this analysis, together with the model's promising applications in option pricing, yielded inspiration to create a simple option pricing framework which combines HAR-RV predictions with Monte Carlo simulations. Two option pricing models are proposed: one with jump component, the HAR-BPV option pricing model, and one without, the HAR-RV option pricing model. The two models are assessed against competitive benchmarks: the PBS model and the GARCH option pricing model of Heston and Nandi (2000). The empirical analysis consists of pricing OTM S&P 500 index options in two years with totally different volatility profiles: 2018 and 2022, corresponding to 320,597 and 517,110 observations, respectively. Due to their large increase in popularity in the past decade, particular attention is paid to options with less than two weeks of time to maturity.

In the year 2018, it is found that the HAR-RV option pricing model attains higher accuracy than its jump counterpart for deep-OTM options, whereas the opposite holds for non-extreme moneyness levels. Furthermore, both HAR option pricing models outperform the GARCH and PBS benchmarks for near-ATM and OTM call options. Short-maturity options with non-extreme moneyness levels and deep-OTM options are found to be most accurately priced with the HAR-BPV and GARCH model, respectively.

In the more volatile year 2022, performance of all models generally deteriorates across all categories. Nevertheless, the HAR-RV option pricing model outperforms both its jump extension and all benchmarks for near-ATM and call options, even yielding a lower pricing error for deep-OTM calls than in 2018. Weekly and biweekly call options are priced the best with the HAR-RV option pricing model, while the GARCH model is the most accurate for put options in this maturity spectrum

Although the results are promising, there are some points of attention. First, the simulation approach makes option pricing very straightforward, however this comes at the cost of having to perform large amounts of simulations, especially if jumps are included. Of course, this is not a major issue if only short-term options are considered. Nevertheless, further research could explore the use of variance reduction methods, such as Quasi-Monte Carlo, to decrease the number of required asset path simulations. Second, the HAR option pricing models have difficulties with pricing deep-OTM put options, implying that the left tail behavior of volatility is not properly captured. A possible solution would be to incorporate the leverage effect, which can be done by considering a HAR model with leverage component (Appendix F provides a gentle introduction on how this could be implemented). Finally, the analysis of the HAR option pricing models' performances during a more volatile period shows clear deterioration for most categories of moneyness and maturity. Further research can be done on how to mitigate the sensitivity of the models to extreme events. This could, for example, include different measures for realized volatility and the bipower variation.

References

Alitab, D., Bormetti, G., Corsi, F. & Majewski, A. A. (2019). A realized volatility approach to option pricing with continuous and jump variance components. *Decisions in Economics and Finance*, *42*(2), 639–664.

Andersen, T. G., Bollerslev, T. & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics*, *89*(4), 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Andersen, T. G., Fusari, N. & Todorov, V. (2017). Short-term market risks implied by weekly options. *The Journal of Finance*, *72*(3), 1335–1386.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2009). *Realized kernels in practice: Trades and quotes.* Oxford University Press Oxford, UK.

Barndorff-Nielsen, O. E. & Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, *17*(5), 457–477.

Barndorff-Nielsen, O. E. & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics*, *2*(1), 1–37.

Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, *81*(3), 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Boyle, P., Broadie, M. & Glasserman, P. (1997). Monte carlo methods for security pricing. *Journal of economic dynamics and control*, *21*(8-9), 1267–1321.

Butcher, J. C. (2016). *Numerical methods for ordinary differential equations.* John Wiley & Sons.

Christoffersen, P., Feunou, B., Jacobs, K. & Meddahi, N. (2014). The economic value of realized volatility: Using high-frequency returns for option valuation. *Journal of Financial and Quantitative Analysis*, *49*(3), 663–697.

Christoffersen, P. & Jacobs, K. (2004). The importance of the loss function in option valuation. *Journal of Financial Economics*, *72*(2), 291–318.

Christoffersen, P., Jacobs, K., Ornthanalai, C. & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, *90*(3), 272–297.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Corsi, F., Fusari, N. & La Vecchia, D. (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics*, *107*(2), 284–304.

Corsi, F. & Reno, R. (2009). Har volatility modelling with heterogeneous leverage and jumps. *Available at SSRN*, *1316953*.

Diebold, F. X. & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, *20*(1), 134–144.

Dumas, B., Fleming, J. & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, *53*(6), 2059–2106.

Dupire, B. et al. (1994). Pricing with a smile. *Risk*, *7*(1), 18–20.

Elton, E. J., Gruber, M. J., Comer, G. & Li, K. (2005). Spiders: Where are the bugs? In *Exchange traded funds: Structure, regulation and application of a new fund class* (pp. 37–59). Springer.

Engle, R. F. & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, *48*(5), 1749–1778.

Geweke, J. & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of time series analysis*, *4*(4), 221–238.

Granger, C. W. & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis*, *1*(1), 15–29.

Hansen, P. R., Lunde, A. & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497.

Harvey, D., Leybourne, S. & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, *6*(2), 327–343.

Heston, S. L. & Nandi, S. (2000). A closed-form garch option valuation model. *The review of financial studies*, *13*(3), 585–625.

Huang, X. & Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of financial econometrics*, *3*(4), 456–499.

Hull, J. & White, A. (1987). The pricing of options on assets with stochastic volatilities. *The journal of finance*, *42*(2), 281–300.

Jacquier, E., Polson, N. G. & Rossi, P. E. (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, *20*(1), 69–87.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management science*, *48*(8), 1086–1101.

Melino, A. & Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of econometrics*, *45*(1-2), 239–265.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, *3*(1-2), 125–144.

Mincer, J. A. & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). NBER.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B. & Ward, J. R. (2008). *Fractals and intrinsic time-a challenge to econometricians.* SSRN.

Oosterlee, C. W. & Grzelak, L. A. (2019). *Mathematical modeling and computation in finance: with exercises and python and matlab computer codes.* World Scientific.

Russell, J. R. & Bandi, F. M. (2004). Microstructure noise, realized volatility, and optimal sampling. In *Econometric society 2004 latin american meetings.*

VettaFi. (2024). *Most popular etfs: Top 100 etfs by trading volume.* Retrieved 2024-05-11, from

https://etfdb.com/compare/volume/

WRDS. (2024). *Data products and vendors at wrds.* Retrieved 2024-05-11, from https://wrds-www.wharton.upenn.edu/

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, *100*(472), 1394–1411.

A HAR-RV Performance in Years of Turmoil

Consider now the full sample running from January 2, 2013, up until December 30, 2022, that is, including the periods of high turmoil in the years 2020 up until 2022. In Table 9, the forecasting results are shown where as before, RMSE is the Root Mean Squared Error, MAE is the Mean Absolute Error, and R^2 denotes the coefficient of determination of the Mincer-Zarnowitz (MZ) regression. It is clear that, compared with the results in Section 3, performance of all models deteriorates by a large amount, resulting in the RMSE and MAE being three and two times as large, respectively, for all models. Furthermore, the forecasts are less efficient, as can be seen from the MZ regression R^2, which is around 20 percentage points smaller for all models and forecasting horizons. This is supported by the rejection of the joint null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ for all models and horizons in Table 10.

In terms of relative performance, the HAR-RV model still outperforms all benchmarks, with the exception of the AR(3) and AR(22) models for a forecasting horizon of two weeks. When looking at the (modified) Diebold-Mariano test statistic p-values in Table 11, it is visible that the HAR-RV model is not significantly better than the benchmarks over all forecasting horizons at the 5% level, with the exception of the AR(22) model for one-day-ahead forecasts.[20]

In all, the HAR-RV model seems to struggle to perform high-quality forecasts in periods of turmoil, likely due to the large amount of jumps in the asset returns. Although one can try to incorporate a jump component in the HAR-RV model, resulting in a HAR-J-RV model in the spirit of Andersen, Bollerslev and Diebold (2007), jumps are not predictable. One therefore needs to include a form of "expected risk" in the model. Implied volatility is known to contain information on trader expectation of future volatility and hence, this could be incorporated in the HAR-RV model to potentially reduce the impact of jumps (that is, new market information) on the model's performance.

Table 9: Out-of-sample Forecasting Performance.

	$h = 1$			$h = 5$			$h = 10$		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
AR(1)	0.096	0.033	0.373	0.448	0.176	0.295	1.180	0.395	0.130
AR(3)	0.092	0.032	0.422	0.429	0.160	0.412	1.170	0.356	0.236
AR(22)	0.097	0.035	0.399	0.452	0.173	0.416	1.115	0.377	0.288
ARFIMA	0.147	0.109	0.255	0.680	0.547	0.236	1.414	1.109	0.123
HAR	0.090	0.032	0.438	0.418	0.158	0.440	1.171	0.348	0.265

Notes: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the R^2 of the MZ regressions of the AR(1), AR(3), ARFIMA$(5, d, 0)$, and HAR-RV models for the SPDR ETF over the full sample (January 2, 2013, up until December 30, 2022). Forecasting horizons of one day, one week, and two weeks are considered ($h = 1$, 5, and 10, respectively). The AR(1), AR(3), and HAR-RV models are re-estimated daily using a moving window of 1000 observations. This is also the case for the ARFIMA model, the only difference being that prior to the forecasting exercise, the parameter d is estimated on the full sample and used to take the fractional difference of the realized volatility series. Forecasts are then obtained using an AR(5) model on the fractionally differenced series.

[20] Performing the MCS procedure of Hansen et al. (2011) at 5% significance and with squared loss results in all models, with the exception of ARFIMA, being included in the MCS, for all forecasting horizons. This matches with the findings of the DM test.

Table 10: MZ regression Wald test p-values.

	$h=1$	$h=5$	$h=10$
AR(1)	0.000	0.000	0.000
AR(3)	0.000	0.000	0.000
AR(22)	0.000	0.000	0.000
ARFIMA	0.000	0.000	0.000
HAR	0.000	0.000	0.000

Notes: p-values of the Wald test statistics for the joint null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ in the MZ regressions of Table 9.

Table 11: DM test p-values.

	HAR		
	$h=1$	$h=5$	$h=10$
AR(1)	0.060	0.175	0.451
AR(3)	0.071	0.080	-
AR(22)	0.016	0.086	-
ARFIMA	0.000	0.000	0.159

Notes: DM test p-values for the null of equal forecast accuracy (accuracies in Table 9) between the HAR-RV model and the model in the corresponding row, for a given forecasting horizon. The alternative is one sided (that is, the MSE of the HAR-RV model is smaller than the MSE of the other competing model. The p-values correspond to the modified DM statistics of Harvey et al. (1997).

B HAR-RV Simulation Study

This section gives a brief illustration on the capability of the HAR-RV model to reproduce stylized facts of financial returns and volatility series. This is done by simulating a financial returns and volatility series, and comparing it with the empirical SPDR ETF data described in Section 3. Following Corsi (2009), the HAR-RV process is simulated at a frequency of two hours (i.e., $M = 12$ in equation (3)). The simulated dynamics are then

$$\begin{aligned} r_t^{(2h)} &= \sigma_t^{(d)} \varepsilon_t, \\ \sigma_{t+2h}^{(d)} &= c + \beta^{(d)} RV_t + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \omega_{t+2h}^{(d)}, \end{aligned} \tag{27}$$

where $r_t^{(2h)}$ represents the two-hour-frequency returns on day t and $\sigma_t^{(d)}$ is the daily "latent" volatility. The innovations ε_t are assumed to be standard normally distributed, and the coefficients c, $\beta^{(d)}$, $\beta^{(w)}$ and $\beta^{(m)}$ are set to 0.015, 0.25, 0.5 and 0.1, respectively. These values are chosen based on the fitted HAR-RV model on the full sample of the daily SPDR ETF realized volatility, which are 2,518 daily observations running from January 2, 2013, up until December 30, 2022. The simulated series consists of 50,000 daily observations (about 200 years).

Looking at the daily returns series in Figure 3, one would not directly say the series look similar. This is mainly due to the extremely large returns around observation 1800, which corresponds to the COVID-19 crisis of 2020. Apart from this observation, the returns series show similar behavior. The daily realized volatility series in Figure 4 are rather different however: the empirical series shows much more volatility spikes that distinguish themselves from the other observations. This is likely due to asset jumps, which are not included in the simulation, but do occur often in the final 1000 observations of the empirical sample.

Finally, Figure 5 shows the autocorrelation functions of the daily realized volatility series. This shows the desired result: the HAR-RV model reproduces the volatility persistence that is visible in empirical data. Although the HAR-RV model is technically a short-memory process, the aggregation interval used shows a long-memory property similar to the daily SPDR ETF realized volatility series.

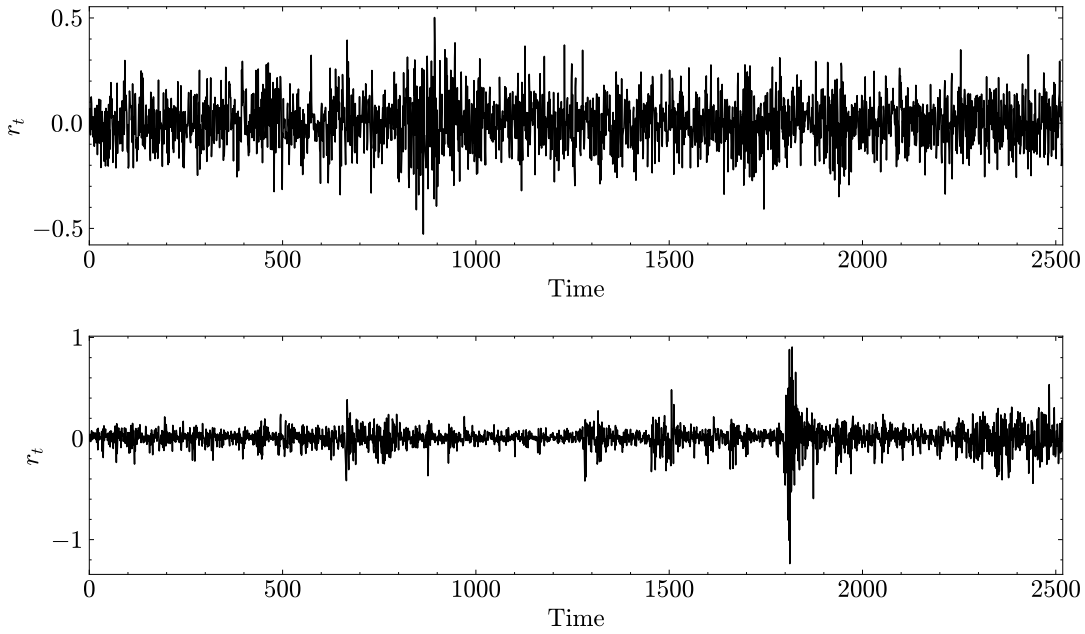


Figure 3: Simulated daily returns series (upper) and empirical SPDR ETF returns (lower).

C Proof of $\mathbb{Q}$-dynamics of Jump-Diffusion Process

Proof. Consider again the SDE in equation (12). Applying Itô's lemma yields

$$\frac{dS(t)}{S(t)} = \left(\mu(t) + \frac{1}{2}\sigma^2(t)\right) dt + \sigma(t)dW(t) + \left(e^J - 1\right) dX_{\mathcal{P}}(t). \tag{28}$$

Using a similar approach as in Oosterlee and Grzelak (2019), define the money-savings account $M(t)$ (also called the "numéraire"). The following holds: $M(0) = 1$, $M(t) = M(s)e^{r(t-s)}$ for $s \leq t$, and $dM(t) = rM(t)dt$. Under $\mathbb{Q}$, the discounted asset price is a martingale. Hence, one should find the condition under which $Y(t) = S(t)/M(t)$ is a martingale. Equivalently, the expected value of the dynamics of $Y(t)$ should be zero. Using the product rule, it holds that

$$dY(t) = \frac{dS(t)}{M(t)} - \frac{rS(t)dt}{M(t)}, \tag{29}$$

and hence the expected value of $dY(t)$ under the physical measure is

$$\mathbb{E}[dY(t)] = \mathbb{E}\left[\mu(t)S(t) + \frac{1}{2}\sigma^2(t)S(t) - rS(t)\right] dt + \mathbb{E}[\sigma(t)S(t)dW(t)] + \mathbb{E}\left[(e^J - 1)S(t)dX_{\mathcal{P}}(t)\right], \tag{30}$$

which should be equal to zero. By using the properties of Brownian motions and Poisson processes, and that these are independent when on the same probability space, one gets

$$\begin{aligned}\mathbb{E}[dY(t)] &= \mathbb{E}\left[\mu(t)S(t) + \frac{1}{2}\sigma^2(t)S(t) - rS(t)\right] dt + \mathbb{E}\left[(e^J - 1)S(t)\right]\xi dt \\ &= \left(\mu(t) - r + \frac{1}{2}\sigma^2(t) + \mathbb{E}\left[\xi(e^J - 1)\right]\right)\mathbb{E}[S(t)]dt = 0,\end{aligned} \tag{31}$$

which only holds when $\mu(t) = r - \frac{1}{2}\sigma^2(t) - \xi\mathbb{E}[e^J - 1]$. Substituting this into equation (28) yields the dynamics for the jump-diffusion process under $\mathbb{Q}$. □

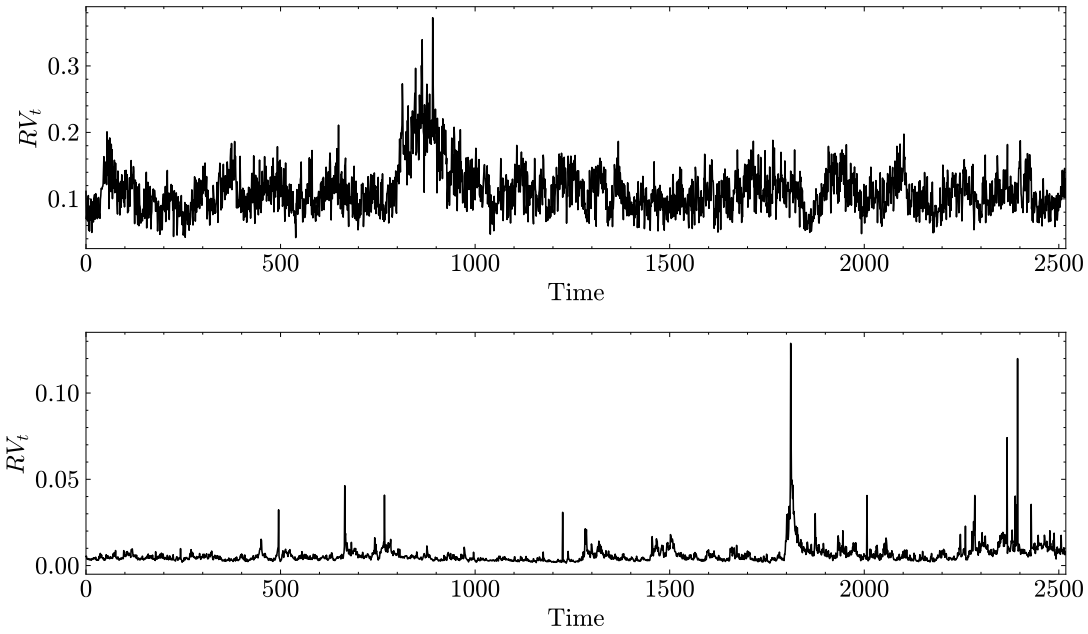


Figure 4: Simulated daily realized volatility series (upper) and empirical SPDR ETF daily realized volatility obtained with the two-scales estimator of Zhang et al. (2005) (lower).

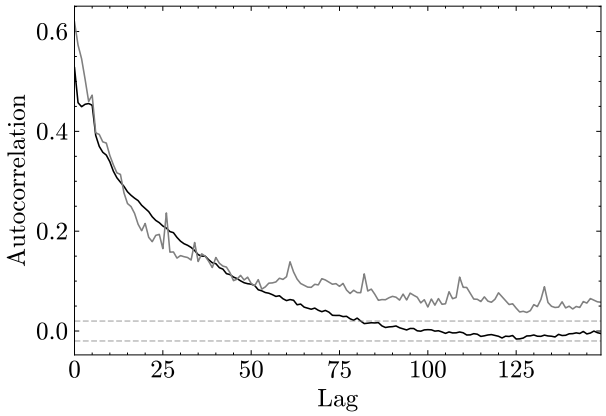


Figure 5: Autocorrelation function of simulated daily realized volatility (black) and empirical SPDR ETF realized volatility (gray). Dotted lines represents 95% confidence bounds.

D Test for Significant Jumps

The statistic of Barndorff-Nielsen and Shephard (2004) shows that under the absence of jumps and for an increasing sampling frequency it holds that

$$\frac{RV_t^2 - BPV_t^2}{\left[\left(\mu_1^{-4} + 2\mu_1^{-2} - 5\right) M^{-1} IQ_t\right]^{\frac{1}{2}}} \xrightarrow{d} N(0,1), \tag{32}$$

where μ_r for $r > 0$ is the r^{th} absolute moment of a standard normal random variable,[21] and $IQ_t = \int_{t-1}^{t} \sigma^4(s)\, ds$ is the so-called "integrated quarcity". A jump-robust estimator for IQ_t (in similar fashion that BPV_t is a jump-robust measure for the integrated variance) is the "tri-power quarcity", equal to

[21] More precisely, $\mu_r = 2^{\frac{r}{2}} \frac{\Gamma((r+1)/2)}{\Gamma(1/2)}$, where $\Gamma(\cdot)$ is the gamma function.

$$TPQ_t = \mu_{\frac{4}{3}}^{-3} \frac{1}{M} \sum_{i=3}^{M} \left|r_{t_{i-2}}\right|^{\frac{4}{3}} \left|r_{t_{i-1}}\right|^{\frac{4}{3}} \left|r_{t_i}\right|^{\frac{4}{3}}, \tag{33}$$

which converges in probability to IQ_t.[22] The test statistic of Barndorff-Nielsen and Shephard (2004) is then obtained by replacing IQ_t by TPQ_t is equation (32). Huang and Tauchen (2005) however states that this statistic tends to indicate that there is a jump when there is none. This thesis thus uses a statistic that yields better finite sample performance, given by

$$\frac{\left(RV_t^2 - BPV_t^2\right)/RV_t^2}{\left[\left(\mu_1^{-4} + 2\mu_1^{-2} - 5\right) M^{-1} \max(1, TPQ_t/BPV_t^4)\right]^{\frac{1}{2}}}. \tag{34}$$

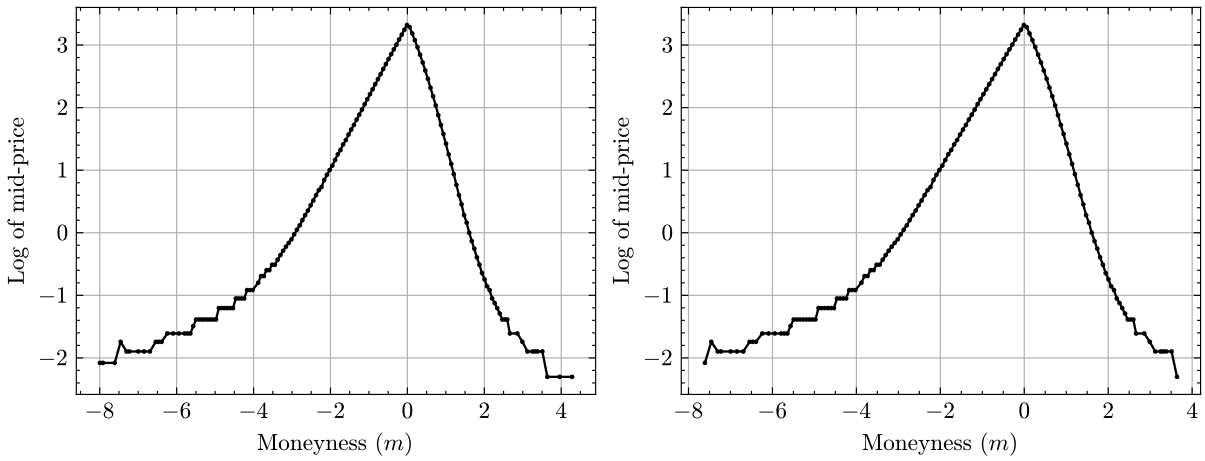
Finally, the test involves choosing a significance level (often small, such as 1% or 0.1%) and performing a right-tailed Z-test (as $RV_t^2 - BPV_t^2$ is truncated to be positive).

E Option Data Cleaning

The first set of filters is similar to Andersen et al. (2017) and are as follows: (i) For each tenor, there are at least 10 distinct option quotes across all strikes; (ii) The maturity is less than or equal to 180 calendar days; (iii) Moneyness (as defined in Section 5) not extreme ($-15 \leq m \leq 5$); (iv) The ratio of the best bid and best offer is less than five (and hence the best bid must be strictly positive). The first filter mitigates noise by diversifying measurement errors across multiple contracts at each maturity. The removal of long-term contracts is common practice. The omission of both extreme moneyness options and options with large spreads ensures that the options are liquid enough (and hence fairly priced).

Options with nine calendar days or less to maturity obtain some additional filters. This because these weekly options are cheaper than biweekly or longer-dated options, and due to the relatively large tick size of \$0.05, the relative error by the rounding of option prices is much larger. The additional filters (not all equivalent to Andersen et al. (2017)) are as follows: (i) Moneyness is restricted to $-8 \leq m \leq 5$; (ii) The last traded contract of an option must be on the day of observation itself; (iii) a final condition to remove stale quotes. This final condition addresses the fact that deep OTM options have a tick size close to the actual value of the option and are the least liquid of all options considered. As a result, identical quotes can be found in the extreme moneyness spectrum. These options are removed as follows: for OTM put options, all quotes at the end of the moneyness spectrum are removed until the mid-price (that is, the middle of the best bid and offer) is smaller than all other quote mid-prices that are closer to the moneyness of zero. For OTM call options, the same is done but in the opposite direction. Figure 6 demonstrates this final procedure for an arbitrary observation day and tenor in the sample.

[22] The tri-power quarcity is calculated using the same sampling frequency of returns as for the other realized measures in this thesis (i.e., ten tick returns).



(a) Prior to final filter

(b) After final filter

Figure 6: Illustration of the additional data-filter applied to weekly options (maturity of 9 days or lower) for an arbitrary observation day and tenor. The observation day is January 12, 2022 and the expiration is January 18, 2022 ($\tau = 6$), which corresponds to 146 contracts satisfying this criterion prior to filtering (6a). After filtering (6b), 4 contracts have been removed: 2 OTM put (left tail), and 2 OTM call (right tail).

F Extension: The Leverage Effect

From Section 5, it became clear that the option pricing models proposed by this thesis fail to capture the left tail of the volatility distribution. A possible explanation for this is that they do not take into account the leverage effect: returns are negatively correlated with volatility. This section demonstrates its inclusion in the HAR-RV option pricing model (the HAR-BPV case is analogous).

As before, a HAR-RV model is estimated using past data. The difference being that the model now takes into account the leverage effect as in Corsi and Reno (2009). That is, if r_t is the return on day t, the average return over the previous week and month are calculated as

$$r_t^{(w)} = \frac{1}{5}\sum_{i=0}^{4} r_{t-i} \quad \text{and} \quad r_t^{(m)} = \frac{1}{22}\sum_{i=0}^{21} r_{t-i}, \tag{35}$$

respectively. The HAR-RV model with leverage is defined as

$$\begin{aligned} RV_{t+1} = c &+ \beta^{(d)} RV_t + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} \\ &+ \gamma^{(d)} r_t^{-} + \gamma^{(w)} r_t^{(w)-} + \gamma^{(m)} r_t^{(m)-} + \omega_{t+1}, \end{aligned} \tag{36}$$

where $r_t^- = \min(r_t, 0)$, $r_t^{(w)-} = \min(r_t^{(w)}, 0)$ and $r_t^{(w)-} = \min(r_t^{(m)}, 0)$. The h-day-ahead forecasts can again be obtained iteratively, which is only possible since the asset path is simulated until the time to expiration. Note that this means that for each asset path, a different realized volatility forecast series is obtained (as each path yields a different returns series). This results in a computationally expensive option pricing method and is therefore left as a topic for future research, in combination with more efficient simulation strategies.

G Brief Description of Code

As requested, an additional ZIP-file containing all programming is handed in. This section provides a brief description of each file. Code has been written in both Python and R. The programming code contains additional details in the form of comments.

- `get_returnsdata.ipynb` is used to obtain the Trade and Quote (TAQ) data of the NYSE via WRDS (2024) on the S&P 500 SPDR ETF, and calculates the realized measures used in this thesis. Data cleaning is also performed in this file.
- `get_optiondata.ipynb` is used to obtain data from OptionMetrics via WRDS (2024) on options written on the S&P 500 index. The data is also cleaned in this file.
- `har.ipynb` prepares the data, fits the HAR model, and performs the forecasting exercises.
- `simulation.ipynb` performs the simulation study on the HAR-RV model and creates images.
- `optiondfs.ipynb` splits the option data in the previously discussed categories of moneyness and maturity.
- `jump.ipynb` analysis of jump behavior, and estimation of parameters of nonsymmetric double exponential distribution for the jump magnitude.
- `garch.ipynb` estimation of GARCH benchmark model parameters using iterative MLE.
- `gbm.ipynb` performs the option pricing routine for all categories. This file yields most of the results in Section 5. Inspiration for this code has partially been found in Oosterlee and Grzelak (2019), which provides an example Python code for Euler discretization of a Geometric Brownian Motion and the simulation of it.
- `ARFIMA.R` estimation of ARFIMA parameter d and fractional integration of realized volatility series.
- `DM_MCS.R` calculation of (modified) Diebold-Mariano test for equal predictive accuracy, as well as the Model Confidence Set procedure.