

M2 Thesis

Portfolio Optimization with Options

ANDREW BOOMER

October 6, 2021

Table of Contents

1	Introduction	3
2	The Model	6
	2.1 Portfolio Optimization Method	6
	2.2 Returns	8
	2.3 Portfolio Optimization Steps	9
3	Data	12
4	Estimation	18
5	Results	20
	5.1 Portfolio Optimization: Initial Model	20
	5.2 Portfolio Optimization: Cutoff Adjusted	24
6	Conclusion	28
\mathbf{A}	Options Overview	32
в	Detailed ExPost Returns	34

Abstract

In this paper, I extend the portfolio optimization with options approach layed out in Faias and Santa-Clara (2017). I do this by replacing their use of an adhoc procedure to estimate realized volatility by introducing a rolling GARCH(1, 1) model to estimate the volatility of log returns. Additionally, this optimization is done on a more recent set of options data, which includes the COVID pandemic. Contrary to Faias and Santa-Clara (2017), I find a lower sharpe ratio from the options optimization compared to the S&P500 in the updated time period. I find that when the size of the optimized weights are not cutoff, the model takes an extreme short position in ATM calls of -23.4% in response to the Covid shock, which partially mitigates the loss from the bear market.

1 Introduction

Portfolio optimization, or portfolio selection, contains theories and techniques from a broad range of subjects. It is an important area of research for both academics and investors, and can provide a mathematical grounding for trading decisions while reducing risk exposure. A portfolio optimization model can also be the foundation for an automated trading software program, which Malmgren and Stys (2011) notes accounted for an average of 60%, and sometimes up to 80%, of equity trading volume on the flash crash of May 6, 2010. Developing an automated options trading program that can be deployed for actual retail trading is, in fact, the initial motivation of this paper. Portfolio selection is the theoretical economic foundation upon which an automated trading program is built, so this paper represents the first step in the trading programs' development. Further considerations and extensions to transition this model to application in real world trading are discussed in Section 6.

Portfolio selection was popularized with the seminal paper by Markowitz (1952). That paper applied a mean-variance framework to optimize the set of weights of a portfolio of stocks. Since then, there have been many extensions to this portfolio selection approach. This paper extends the optimization on a basket of stocks to a basket of options. For an overview on the basics of options, see Appendix A. While portfolio selection with options has traditionally been less common in the literature, there are good reasons to include them. Zhao and Palomar (2018) discuss two, (1) while under the scrict Black-Scholes assumption of no riskless arbitrage, returns from any option portfolio can be replicated with a portfolio of stocks, in the real world, including options can achieve excess returns; (2) options allow the investor to make directional bets on stock price movements, limiting certain risk exposure. While portfolio selection with options is not as common, Zhao and Palomar (2018) utilize the mean-variance framework as in Markowitz (1952), and add options into the optimization problem. They use the Black-Scholes assumption of stock returns following a geometric brownian motion to optimize the set of weights. This paper, however, most closely follows the work in Faias and Santa-Clara (2017), who introduce an expected utility function in place of the mean-variance optimization. They argue, for three main reasons, that the mean-variance framework is not well suited to options. First, option returns are not effectively described by their first and second moments. Second, large amounts of historical options data is difficult to acquire, so distributional estimation is imprecise. Third, the bid-ask spread introduces transactions costs which are hard to account for in the mean-variance framework.

The last consideration of the portfolio selection problem, specifically in the context including option contracts, is whether the options are held until expiration or the portfolio is dynamically rebalanced. Zhao and Palomar (2018) directly extend the dynmically rebalanced stock portfolio of Markowitz (1952) by using the Black-Scholes option pricing model to analytically formulate the change in the value of a portfolio containing a basket of stocks and option contracts. Following Faias and Santa-Clara (2017), I conduct the portfolio selection problem with options assuming the options are held until expiration. While simpler to implement, this method does have a couple disadvantages. First, holding until expiration relies on long term return estimation, one month in the case of this paper, which can be unreliable. Additionally, holding until expiration does not allow the investor to exit their position if the risk exceeds their tolerance or the market moves in an unfavorable direction. Dynamic rebalancing is further discussed in Section 6.

An overview of the optimization assumptions applied in this context provides some foundation for the portfolio selection approach in this paper, as well as context for future research and extensions. Hannah (2015) provides an excellent summary of the data settings for an optimization problem. Additionally, the other important consideration in structuring an optimization model is whether it is possible to formulate as convex.

For some objective function to be minimized, $\mathcal{F}(x,\xi)$, where $\xi \in \Xi$ is a random variable and $x \in X$ is the choice variable, I will assume that (1) ξ is exogenous of x such that \hat{x} does not affect the distribution of ξ ; (2) The data generation is constructive such that given ξ , $\mathcal{F}(x,\xi)$ can be calculated for all $x \in X$; (3) The data is observational such that data are generated *a priori* and new data cannot be generated. While these assumptions are standard in the portfolio selection literature, they may not always hold. For example, investment firms trading sufficiently large volumes are able to influence the future price of an asset. As described in Gsell (2008), institutional investors use algorithms to slice large orders into smaller ones to avoid adverse price movements through exhausting the current liquidity in the market. This may even be a larger issue for options than for stocks, since the liquidity of an option contract can vary significantly depending on its moneyness. In this paper, only S&P500, a highly liquid asset, is considered, and potentially illiquid far out of the money contracts are also not considered. However, generalizing the model to include other assets or further out of the money options could lead to low liquidity environments. In this case, assumption (1) may not hold.

Additionally, portfolio selection is not necessarily a convex optimization problem, it must be specifically formulated to be convex. Markowitz (1952), Zhao and Palomar (2018), and Faias and Santa-Clara (2017) assign real number weights to the set of options (stocks) as the decision variables in the optimization. This makes the feasible set of the optimization convex, and given the convex objective functions used, the entire problem by extension. Convex optimization is more tractable and less computationally expensive than nonconvex, so the convex formulation is widely used in the literature. Non-convex extensions are discussed further in Section 6.

In Section 2.1 I first provide an overview of the portfolio selection method, where I define the objective function to be optimized, as well as how the returns are constructed. In Section 2.2, I introduce the GARCH(1, 1) volatility that is used to model the volatility of the underlying S&P500 asset. In Section 2.3, the detailed steps of the portfolio optimization problem are presented. I show how optimizing for next period's wealth can be reformulated in terms of the option returns in the portfolio, accounting for the inclusion of the risk-free asset. I then detail the realized returns in terms of the optimized weights which are used to evaluate the performance of the model. In Section 3, I describe the data sources for the options, risk-free asset, and underyling asset. I then describe the multi-step data filtering process for preparing the options data. I show various summary statistics and visualizations of each of the datasets, and comment on the historical behavior of each. In Section 4, I present the first period estimation of the rolling GARCH(1, 1) model, commenting briefly on the estimated parameters. In Section 5, I present the additional variation of the model including the introduction of a cutoff to ensure adequate shrinkage of the weights. I show that the optimization method performs worse than the S&P500. I present detailed analyses of the optimized weights and returns, showing that the model is generally net short on options, and long on the risk-free asset. I compare the performance of the basic model with weight cutoff variation, and show that the restriction induces worse performance. In Section 6, I conclude, review the model and findings, and discuss areas for further research.

2 The Model

2.1 Portfolio Optimization Method

I now introduce the 6-step portfolio selection method, modeled after Faias and Santa-Clara (2017). In this method, I use a risk free asset, modeled by the 1-month treasury bill, and a risky asset underlying the call and put options, modeled by the S&P500. The portfolio selection problem optimizes a set of weights among the risk free asset and call/put options. Formally, denoting A_t as the wealth at time t, I solve:

$$\max_{\mathbf{W}_t} E[U(A_{t+1})|F_t]$$

Where F_t is the available information at time t, E is the expectation operator, and U is the utility function specified later. \mathbf{W}_t is the vector of weights allocated to each asset.

Additionally, I define S_t as the price of the underlying risky asset at time t, rf_t as the risk free rate at time t, $C_{t,k}$ $(P_{t,k})$ as the price of the call (put) option with strike price k

at time t which expires at time t + 1. Lastly, the continuously compounded returns of S_t are defined as $y_t = log(S_t) - log(S_{t-1})$

Suppose that at time t, there are C number of call option contracts with strike prices $K_{t,c_1}, K_{t,c_2}, \dots, K_{t,c_C}$, and P number of put option contracts with strike prices $K_{t,p_1}, K_{t,p_2}, \dots, K_{t,p_P}$. Then the wealth at time t + 1 is given by:

$$A_{t+1} = A_t \left(\left(1 - \sum_{i=1}^C \omega_{t,c_i} - \sum_{i=1}^P \omega_{t,p_i}\right) rf_t + \sum_{i=1}^C \omega_{t,c_i} r_{t+1,c}(K_{t,c_i}) + \sum_{i=1}^P \omega_{t,p_i} r_{t+1,p}(K_{t,p_i}) \right)$$

where $r_{t+1,c}(k)$ $(r_{t+1,p}(k))$ denotes the return of the call (put) option with strike price k. Formally:

$$r_{t+1,c}(k) = \frac{\max\{S_{t+1} - k, 0\}}{C_{t,k}} - 1$$
$$r_{t+1,p}(k) = \frac{\max\{k - S_{t+1}, 0\}}{P_{t,k}} - 1$$

Finally, the portfolio returns, rp_{t+1} (i.e. $rp_{t+1} = \frac{A_{t+1}}{A_t} - 1$) are defined as:

$$rp_{t+1} = (1 - \sum_{i=1}^{C} \omega_{t,c_i} - \sum_{i=1}^{P} \omega_{t,p_i})rf_t + \sum_{i=1}^{C} \omega_{t,c_i}r_{t+1,c}(K_{t,c_i}) + \sum_{i=1}^{P} \omega_{t,p_i}r_{t+1,p}(K_{t,p_i})$$

Transaction costs are modeled through the bid-ask spread, and are incorporated through duplicating the basket of option contracts into long and short side options. Long options enter at the end of day ask price, and short options enter at the end of day bid price. Incorporating the bid-ask spread in this way converts this model to a constrainted optimization problem, with a no short-selling constraint such that $\mathbf{W}_t > 0$. Short side option returns enter the optimization model multiplied by -1. This is an important consideration, and Santa-Clara and Saretto (2009) find that transaction costs and margin calls can explain some of the pricing anomolies seen in the options market.

In the framework of this model, since S_t is a random variable conditional on F_t , then rp_{t+1} is also conditionally random. Additionally, I rely on a simulation by replicating

 ${rp_{t+1|t}^n}_{n=1}^N$ conditional on F_t , and transform it into ${A_{t+1|t}^n}_{n=1}^N$. I then solve the following maximization problem on the simulated next period wealth:

$$\max_{\mathbf{W}_t} \frac{1}{N} \sum_{n=1}^N U(A_{t+1|t}^n)$$

which is described in detailed steps in Subsection 2.3. First, the GARCH volatility model is introduced.

2.2 Returns

In this paper I assume that the log asset returns (y_t) follow a normal distribution with a time varying volatility term. Where Faias and Santa-Clara (2017) employ an ad-hoc estimation of realized volatility, I consider a GARCH(1, 1) specification. Therefore, the full model specification for the log returns of the underlying asset is given by the GARCH (1, 1) with zero conditional mean:

$$y_t = \sigma_t \epsilon_t \quad \epsilon_t \sim iid.N(0,1)$$
$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$
$$\omega > 0 \quad \alpha, \beta \ge 0$$

As explained in further detail in Section 3, monthly price data for the S&P 500 is obtained from the Case-Shiller index. This is used to estimate the GARCH(1, 1) model via maximum likelihood. This estimation is done via a rolling estimation scheme where the GARCH parameters are re-estimated at each time period given the newly available data. The option data is available between January 2019 and May 2021, so volatility forecasts are computed for this time period from the rolling GARCH model. Additionally, the volatility forecast at time t + h uses the information of the actual y_t at time t + h - 1. For each time period t + h beginning in January 2019, the variance forecast is:

$$\hat{\sigma}_{t+h}^2 = \hat{\omega} + \hat{\alpha} y_{t+h-1}^2 + \hat{\beta} \hat{\sigma}_{t+h-1}^2$$

The volatility is then $\hat{\sigma}_{t+h} = \sqrt{\hat{\sigma}_{t+h}^2}$

2.3 Portfolio Optimization Steps

• Step 1: Simulate log returns of the underlying asset

To simulate future log returns, the ϵ_t are simulated $(\tilde{\epsilon}_t)$ using a bootstrapping method from historical $\hat{\epsilon}_t$, where $\hat{\epsilon}_t = \frac{y_t}{\hat{\sigma}_t}$. This follows the filtered historical simulation method in Barone-Adesi et al. (2008). At each time t, N samples are chosen with replacement from the historical $\hat{\epsilon}_t$ series, time t inclusive. For a timespan of T months, this gives a bootsrapped matrix of $\tilde{\epsilon}_t$ of size $T \times N$. The volatility forecast time series is then multiplied by each of the N simulation for each of the time periods such that $\tilde{y}_{t+h}^n =$ $\hat{\sigma}_{t+h}\tilde{\epsilon}_{t+h}^n \quad \forall n \in 1, \dots, N$.

• Step 2: Simulate next period prices of the underlying asset

For each period t + h, and for each simulation $n \in N$, the simulated price is:

$$S_{t+h|t+h-1}^n = S_{t+h-1}e^{y_{t+h}^n}$$

This yields a matrix of simulated asset prices with the same shape as the \tilde{y}_t matrix.

• Step 3: Simulate option returns

The option returns are calculated assuming the contract is held until expiration. This

means that the simulated payoff of a call and a put, $\forall n \in N$, respectively are:

$$C_{t+1|t}^{n}(k) = \max(S_{t+1|t}^{n} - k, 0) \qquad k = K_{t,c_{1}}, \cdots, K_{t,c_{C}}$$

$$P_{t+1|t}^{n}(k) = \max(k - S_{t+1|t}^{n}, 0) \qquad k = K_{t,p_{1}}, \cdots, K_{t,p_{F}}$$

$$r_{t+1|t,c}^{n}(k) = \frac{C_{t+1|t}^{n}(k)}{C_{t,k}} - 1 \qquad k = K_{t,c_{1}}, \cdots, K_{t,c_{C}}$$

$$r_{t+1|t,p}^{n}(k) = \frac{P_{t+1|t}^{n}(k)}{P_{t,k}} - 1 \qquad k = K_{t,p_{1}}, \cdots, K_{t,p_{F}}$$

• Step 4: Simulate Portfolio Returns

Let $rp_{t+1|t}^n$ denote the portfolio returns for a simulation n:

$$rp_{t+1|t}^{n} = rf_{t} + \sum_{i=1}^{C} \omega_{t,c_{i}}(r_{t+1|t,c}^{n}(K_{t,c_{i}}) - rf_{t}) + \sum_{j=1}^{P} \omega_{t,p_{j}}(r_{t+1|t,p}^{n}(K_{t,p_{j}}) - rf_{t})$$

This can be vectorized be concatenating the call/put returns and weights.

$$rp_{t+1|t}^{n}(\mathbf{W}_{t}) = rf_{t} + \langle \mathbf{W}_{t}, \mathbf{R}_{t+1|t}^{n} \rangle$$
$$\mathbf{W}_{t} = \left[\omega_{t,c_{1}}, \cdots, \omega_{t,c_{C}}, \omega_{t,p_{1}}, \cdots, \omega_{t,p_{P}} \right]$$
$$\mathbf{R}_{t+1|t}^{n} = \left[r_{t+1|t,c_{1}}^{n}, \cdots, r_{t+1|t,c_{C}}^{n}, r_{t+1|t,p_{1}}^{n}, \cdots, r_{t+1|t,p_{P}}^{n} \right] - \left[rf_{t}, \cdots, rf_{t} \right]$$

• Step 5: Present utility function

I use a power utility function which has the property of constant relative risk aversion (CRRA). Bliss and Panigirtzoglou (2004) estimate the CRRA parameter γ as 4 for the S&P500 index option data, but I follow the use of $\gamma = 10$ by Faias and Santa-Clara (2017). Using a higher gamma parameter that was estimated has a similar effect as shrinking the portfolio weights, which is employed in practice as extreme portfolio weights would lead

to increased risk exposure.

$$U(A) = \begin{cases} \frac{1}{1-\gamma} A^{1-\gamma}, & \text{if } \gamma \neq 1\\ \log(A), & \text{if } \gamma = 1 \end{cases}$$

In terms of the vectorized implementation, across all N simulations, I maximize the mean utility of next periods' simulated wealth:

$$\max_{\mathbf{W}_{t}} \sum_{n=1}^{N} U(A_{t+1|t}^{n}) = \max_{\mathbf{W}_{t}} \sum_{n=1}^{N} U(A_{t}(1+rp_{t+1|t}^{n}(\mathbf{W}_{t})))$$

Since A_t is known at time t and is independent of the decision variable \mathbf{W}_t , this optimization problem is equivalent to optimizing the simulated portfolio returns:

$$\max_{\mathbf{W}_t} \sum_{n=1}^N U(1 + rp_{t+1|t}^n(\mathbf{W}_t))$$

• Step 6: Calculate Ex-Post returns

The vector of portfolio weights that solves the maximization problem in Step 5 can be written as:

$$\mathbf{W}_t^* = \arg \max_{\mathbf{W}_t} \sum_{n=1}^N U(1 + rp_{t+1|t}^n(\mathbf{W}_t))$$

The realized ex-post returns are calculated with a similar procedure as the simulated returns. The realized underlying asset price at t + 1, S_{t+1} is used in place of $S_{t+1|t}^n$.

$$C_{t+1|t}(k) = \max(S_{t+1} - k, 0) \qquad k = K_{t,c_1}, \cdots, K_{t,c_C}$$

$$P_{t+1|t}(k) = \max(k - S_{t+1}, 0) \qquad k = K_{t,p_1}, \cdots, K_{t,p_P}$$

$$r_{t+1|t,c}(k) = \frac{C_{t+1|t}(k)}{C_{t,k}} - 1 \qquad k = K_{t,c_1}, \cdots, K_{t,c_C}$$

$$r_{t+1|t,p}(k) = \frac{P_{t+1|t}(k)}{P_{t,k}} - 1 \qquad k = K_{t,p_1}, \cdots, K_{t,p_P}$$

Then vectorizing as before:

$$rp_{t+1|t}(\mathbf{W}_{t}^{*}) = rf_{t} + \langle \mathbf{W}_{t}^{*}, \mathbf{R}_{t+1|t} \rangle$$
$$\mathbf{R}_{t+1|t} = \left[r_{t+1|t,c_{1}}, \cdots, r_{t+1|t,c_{C}}, r_{t+1|t,p_{1}}, \cdots, r_{t+1|t,p_{P}} \right] - \left[rf_{t}, \cdots, rf_{t} \right]$$

Next periods wealth is then:

$$A_{t+1} = A_t (1 + rp_{t+1|t}(\mathbf{W}_t^*))$$

3 Data

The S&P500 monthly data was downloaded from Robert Shiller's website, within the U.S. Stock Markets 1871-Present and CAPE Ratio dataset. This is monthly data spanning from January 1950 to May 2021, including more recent data than Faias and Santa-Clara (2017). The S&P Composite price is used to calculate the log returns (y_t) , which are calculated as $y_t = log(S_t) - log(S_{t-1})$, where S_t is the monthly average of the underlying S&P500 daily closing price.

Table 1 shows these summary statistics for the log returns and standardized returns. While the autocorrelation, Ljung-Box test, and ARCH test statistics all show a meaningful decrease in the standardized returns compared to the log returns, the skew and kurtosis show only a modest reduction (-1.22 to -1.2, and 5.12 to 4.68 respectively).

Table 1:S&P 500 Index Returns Summary Statistics

		Log Returns y	Standardized Returns $\hat{\epsilon}_t$			
Statistics	1950 - 2018	2019 - 2021	1950 - 2021	1950 - 2018	2019 - 2021	1950 - 2021
No. of obs.	828	33	861	828	33	861
Skewness	-1.01	-3.7	-1.22	-0.89	-4.25	-1.2
Excess kurtosis	3.99	15.91	5.12	2.57	19.52	4.68
$ \rho_1(z) $	0.24	-0.02	0.22	0.18	-0.01	0.17
$ \rho_1\left(z^2\right) $	0.15	-0.04	0.11	0.0	-0.06	-0.01
$Q_1(z)$	46.52	0.01	42.55	28.34	0.0	25.32
	[0.0]	[0.9]	[0.0]	[0.0]	[0.96]	[0.0]
ARCH(1)	19.35	0.06	10.74	0.02	0.09	0.15
	[0.0]	[0.81]	[0.0]	[0.9]	[0.76]	[0.7]

Standardized returns, $\hat{\epsilon}_t = \frac{y_t}{\hat{\sigma}_t^{GARCH}}$, are the estimated GARCH conditional volatility series. No. of Observations, Skew, and Kurtosis are presented. Lag 1 autocorrelation, $\rho_1(z)$, and squared autocorrelation, $\rho_1(z^2)$, are additionally shown. $Q_1(z)$ is the Lag 1 LjungBox test of autocorrelation. ARCH(1) is the Lag 1 Engle Test for Autoregressive Conditional Heteroskedasticity (ARCH).

The option data is obtained from the Chicago Board of Exchange (CBOE), and spans from January 2019 to May 2021. Table 2 gives a list of available variables.

Data Descriptions are presented for the relevant variables from the CBOE							
Variable Name	Description						
underlying_symbol	Asset Symbol						
quote_date	Quote Time						
expiration	Contract Expiration Date						
strike	Contract Strike Price						
option_type	Contract Type: Call/Put						
open	Option Price at Open						
close	Option Price at Close						
high	High Option Price						
low	Low Option Price						
Volume	Contract Trading Volume						

Table 2:CBOE Variable Descriptions

Figure 1 shows histograms of the historic log returns and standardized returns for all months prior to January 15, 2019, the start of the CBOE data. Confirming the statistics in Table 1, the distributions of both log and standardized returns show a negative skew, departing from normality.



Figure 1: Historic Log/Standardized Returns

Histograms of historic log returns (y_t) and standardized returns $(\hat{\epsilon}_t)$ are presented.

In order to remove unrealistic data, observations that meet any of these 3 criteria are removed from the CBOE options dataset: (1) zero volume traded; (2) bid price lower than \$0.125, (3) bid price lower than ask price. These 3 filters correspond to the columns: 'Volume', 'bid_eod', and 'ask_eod' respectively. The data is then filtered for expiration dates that land on the 3rd Friday of each month. Then option quote times are filtered for 1 month to maturity, 28/29/30/31 days depending on the month and year. Figure 2 shows the number of call and put contracts by date from the CBOE data after filtering.



Figure 2: Count of Option Contracts

Then I further filter the option contracts used in the portfolio optimization by defining four types of contracts. I define an at-the-money (ATM) call/put and an out-of-the-money (OTM) call/put. ATM calls/puts are found by finding the contract for each date with the lowest bid-ask spread and a moneyness between -1% to 1%. OTM calls/puts are found in the same manner, however with moneyness from -5% to -2% and 2% to 5%respectively. Figure 3 shows a time series of these four contracts option closing prices. While Covid was priced into the OTM/ATM puts in March 2020, it was priced into the ATM call a month earlier in February. This indicates the options market believed the S&P500 would increase from February to March, while in March, after the underlying price had already dropped significantly, the options market believed that the underlying price would continue to drop.

Figure 3: Option Prices Time Series



Time Series of option closing prices by type of contract. Time span is from January 2019 to May 2021. February 2020 is marked as a dotted line.

Table 3 shows summary statistics for the realized returns of these four contracts, as well as for an even weighting between the four, the underlying S&P500, and the risk free asset. Figure 4 shows the histograms of the actual returns of each of the four option contracts, with extreme outliers removed. All four option contracts exhibit a positive skewness, while an ATM call has the lowest skewness and kurtosis of the four contracts. The mean returns of both ATM and OTM puts are larger than ATM calls, due to their large positive tail returns. However their standard deviations lead to lower sharpe ratios than the ATM call. The OTM call has a negative mean return, implying that shorting OTM calls could be a profitable strategy. The positive mean returns of ATM and OTM puts is in contrast to the time period used in Faias and Santa-Clara (2017), which may be due to the small sample size and the outsized impact of the maximum returns in February 2020.

	Table 3	:			
Actual Returns for	Options,	S&P500,	and	\mathbf{Risk}	Free

ation strate s. S&P500 asury bill 1	egy. $\frac{1}{N}$), Risk respecti	shows r Free sho vely. M	eturn fe ow actu ean. Ste	or even al retur d. Min.	weight: ns for Max a	ing on eac underlyin re shown	ch of 1 g asse as per
0	Mean	Std	Min	Max	Skew	Kurtosis	SR
ATM Call	13.5	98.3	-100.0	298.3	0.71	0.49	0.14
ATM Put	19.3	561.8	-100.0	2923.5	5.01	23.42	0.03
OTM Call	-55.2	122.9	-100.0	498.1	3.53	12.62	-0.45
OTM Put	89.1	1018.2	-100.0	5383.4	5.10	24.04	0.09
1/N rule	-19.6	107.8	-100.0	480.8	3.58	14.29	-0.18
S&P 500	1.7	5.1	-21.2	6.1	-3.41	13.10	0.33
Risk Free	0.1	0.1	0.0	0.2	0.23	-1.75	1.01

m11 -1 skС

Figure 4: Actual Returns by Option Type

Histograms of actual returns by option type for a hold until expiration strategy are presented. Returns are percentages, time spans is from January 2019 to May 2021. Returns larger than 1000% are removed for clarity.



To model the risk free asset, the 1-month treasury bill time series is used. The daily data is downloaded from the Fred St. Louis Database (Ticker "TB4WK"). This data is then joined to match up with the quote time of the options in the filtered CBOE data.

This produces a time series of annualized risk free returns, which I convert to monthly returns and take as given in the optimization. Figure 5 shows the time series of the 1-month treasury bill returns. While Figure 5 does not go all the way back to 1996 as in Faias and Santa-Clara (2017), the average risk-free return from 2001 through 2013 is 1.57 times higher than the risk-free return in the red area, from 2019 through May 2021. As discussed in further detail in section 5, this discrepancy may impact a strategy that tends to choose longer positions on the risk free asset.

Figure 5: Risk Free Asset Returns





4 Estimation

The rolling window GARCH(1, 1) model described in section 2.2 is estimated using maximum likelihood at a monthly frequency, spanning from December, 2018 to May, 2021. At each time period t, the data used to estimate the GARCH model spans from January, 1950 to t, inclusive. The estimation results for December, 2018 are shown in Table 4. All three coefficients are significantly different from zero, although ω is 0 out to three decimal points. Additionally, the marginal effect of the squared lag log return (α) is lower than the marginal effect of the squared lag conditional volatility (β).

	coef	std err	\mathbf{t}	$\mathbf{P} \! > \! \mathbf{t} $
ω	0.000	0.000	3.508	0.000
α	0.125	0.038	3.309	0.001
β	0.783	0.037	21.404	0.000

Table 4:GARCH(1, 1) Log Returns Model

GARCH(1, 1) model estimation summary is shown. The model is estimated on log returns from January 1950 through December 2018.

Figure 6 shows a time series plot of the log returns and the conditional volatility using the rolling GARCH procedure described in Section 2.2. The spikes in conditional volatility for the great recession and the Covid pandemic are visible in Figure 6 as well.

Figure 6: GARCH Time Series Plot



Figure 7 shows a time series of the three GARCH(1, 1) parameters estimated through the rolling scheme. There is some variation in the $\hat{\alpha}$ and $\hat{\beta}$ parameter values across the time period, concentrated in the first 12-18 months, while the $\hat{\omega}$ parameter shows little variation. $\hat{\alpha}$ jumped 25% from March to April 2020, while $\hat{\beta}$ showed a negligible decline.

This can be interpreted that after the Covid shock, the impact of the previous log return value on conditional volatility, through $\hat{\alpha}$, became more important.



Figure 7: GARCH Parameters Time Series

5 Results

One additional variation of the portfolio optimization technique laid out in section 2 is now presented. This model introduces a cutoff on the elements of the optimized portfolio weight vector \mathbf{W}_t^* , such that the simulated returns are set to 0 for contracts whose optimized weights are greater than 10% in absolute value, then the optimization is run again. While setting $\gamma = 10$ does induce some weight shrinkage, this variation models a scenario where a trader enforces a cutoff for their single contract exposure.

5.1 Portfolio Optimization: Initial Model

Figure 8 shows a histogram of the monthly Ex-post returns. We can see that there is a left skewness in the distribution, thus suggesting non-normality of the monthly option returns. Additionally, both the highest and lowest monthly returns are centered around the beginning of the Covid pandemic. The lowest monthly return of -24.2%, as shown in Table 5 and in full detail in Table 9 of Appendix B, occurs on the March 2020 expiration, and the highest monthly return of 11.5% occurs the next month in April 2020. The lowest return is larger in absolute value than the highest.



Figure 8: Portfolio Optimization Returns

Table 5 compares the summary of the monthly returns of the option portfolio optimization with the returns of the S&P500. The Ex-post returns have a mean of -0.2% and a standard deviation of 5%, resulting in a lower monthly sharpe ratio than the S&P500, -0.03 vs. 0.33 in the time period of the data. Additionally, the Ex-post returns have a lower absolute skew and kurtosis, implying that they are closer to normality than the S&P500 returns.

Table 5:Portfolio Optimization vs. S&P500 Returns

Summary statistics for t	the optimized port	folio and S&I	P500 returns	are shown. The	he
time period is January 2	019 to May 2021.	Mean, std, m	iin, and max	are percentage	es
	The Sharpe Ratio	(SR) is mon	thly.		

	Mean	Std	Min	Max	Skew	Kurtosis	SR
S&P 500	1.7	5.1	-21.2	6.1	-3.41	13.10	0.33
ExPost	-0.2	5.4	-24.2	11.5	-2.75	11.83	-0.03

Table 6 shows the summary statistics of the weights for each type of option contract in the monthly optimization. Contrary to what would be expected from the actual returns by contract in Table 3 of Section 3, the model is long OTM calls on average, while short on average for ATM calls and OTM puts. Additionally, the range of the weights chosen for the OTM call position, between the minimum and maximum, is much smaller than the other three contracts. Figure 9 shows a time series representation of the weights per contract, as well as the risk-free asset weights. This is also shown in detail in Table 9 of Appendix B. The risk-free asset weights range from 97% to 123% on the March 2020 expiration, where the model exploited the price spike in the ATM call with an extreme short position (-23%). On average, the portfolio optimization model is a net seller of options, and is in a net short position 83% of the months. The model overwhelmingly shorts OTM puts, 93% of the months, with an average short position of -1%. Lastly, OTM calls generally have a higher weight relative to OTM puts, while this is reversed for ATM calls/puts.

Mean,	n, min, and max are presented for the optimized portfolio weights. The period is January 2019 to May 2021. All values are percentages.									
		ATM Call	ATM Put	OTM Call	OTM Put	-				
	Mean	-2.3	0.1	0.4	-1.2	_				
	Minimum	-23.4	-3.1	-1.8	-9.6					
	Maximum	1.8	2.7	1.9	0.2					

Table 6:Portfolio Optimization Weights

Figure 9: Portfolio Optimization Weights by Contract



Optimized weights from portfolio optimization are shown by option type, as well as for the call-put difference and risk-free assset. The time period is January 2019 to May 2021.

Figure 10 shows the cumulative returns from January 2019 to May 2021 for the portfolio optimization model, the S&P500, and the risk-free asset. The initial wealth and February 2020 are marked with dotted lines. Prior to Covid, the cumulative returns of the risk-free asset and the ExPost returns track relatively closely, however, each respond to the Covid shock differently. The risk-free asset levels out, in accordance with Figure 5, while the Expost returns and the S&P500 both experience steep losses, with the S&P500 experiencing a 21% loss. However, the S&P500 bounces back more quickly, and has the highest ending cumulative return of the three, with the Ex-post returns experiencing a small rebound and then a leveling out.

Figure 10: **Cumulative Returns**



5.2Portfolio Optimization: Cutoff Adjusted

2019-10

2020-01

2020-04

2020-07

Date

2020-20

2022-02

2022-04

2022-07

2019-01

120

100

80

2019-04

Figure 11 displays a histogram of the cutoff adjusted Ex-post returns, and Table 7 displays a table comparing these returns to the S&P500 returns. Contrary to what might be expected, comparing Table 7 and Table 5, the largest loss for the cutoff adjusted model is larger in absolute value than the basic model (-28% vs -24%). This can be explained by the removal of the large short position in the ATM call from the basic model. While shorting the high price of the ATM call in February 2020 wasn't enough to offset the loss from shorting the ATM put that month, it did partially mitigate it. Additionally, the skew and kurtosis of the cutoff adjusted model are higher as well, meaning it is further from normality than the basic variation.

Figure 11: Portfolio Optimization Returns: Cutoff Adjusted

Histogram of portfolio optimization returns is shown. The time period is January 2019 to May 2021. Simulated returns for contracts with weights greater than 10% in absolute value are forced to zero, and the optimization is rerun.



Table 7:Portfolio Optimization vs. S&P500 Returns: Cutoff Adjusted

Summary statistics for the optimized portfolio and S&P500 returns are shown.	The
time period is January 2019 to May 2021. Mean, std, min, and max are percenta	ges.
The Sharpe Ratio (SR) is monthly. Simulated returns for contracts with weight	ts
greater than 10% in absolute value are forced to zero, and the optimization is rer	run.

	Mean	Std	Min	Max	Skew	Kurtosis	SR
S&P 500	1.7	5.1	-21.2	6.1	-3.41	13.10	0.33
ExPost	-0.3	6.1	-28.5	11.5	-3.24	13.98	-0.05

Comparing the optimized weights of the cutoff adjusted model in Table 8 with those of the basic model in Table 6, the means are actually fairly similar. Looking at the weights in more detail in Figure 12, we can see the effect of the cutoff adjustment on the risk-free asset weights, which are lower when compared to the basic model in Figure 9. This is especially true in the March 2020 expiration, where the basic model had a more extreme spike (123% vs. 112%) in the optimized weight of the risk-free asset. This is due to the discrepancy in the weights of the ATM call on the March 2020 expiration, which was heavily shorted in the basic model (-23.4% vs. 0%). Finally, since the loss was actually larger in March 2020 in the cutoff adjusted model (-29% vs. -24%), removing the heavily shorted ATM call was not advantageous in this situation. However, favoring the basic model over the cutoff variation due to this result may be a case of overfitting, as reducing an extreme weight may prove profitable in other settings.

Table 8:Portfolio Optimization Weights: Cutoff Adjusted

Mean, min, and max are presented for the optimized portfolio weights. The time period is January 2019 to May 2021. All values are percentages. Simulated returns for contracts with weights greater than 10% in absolute value are forced to zero, and the optimization is rerun.

	ATM Call	ATM Put	OTM Call	OTM Put
Mean	-1.5	0.2	0.3	-1.2
Minimum	-5.3	-3.1	-1.8	-9.6
Maximum	1.8	2.7	1.9	0.2

Figure 12: Portfolio Optimization Weights by Contract: Cutoff Adjusted

Optimized weights from portfolio optimization are shown by option type, as well as for the call-put difference and risk-free assset. The time period is January 2019 to May 2021. Simulated returns for contracts with weights greater than 10% in absolute value are forced to zero, and the optimization is rerun.



Figure 13 displays the cumulative returns for a starting wealth of 100\$ for the cutoff adjusted model. The decreased performance of the cutoff adjusted model is visible, where the more extreme loss in March 2020 is visible, and the ending wealth is lower compared to the basic model (\$85.49 vs. \$90.70).







6 Conclusion

The results of the portfolio optimization model presented in this paper show Ex-post returns which are on average lower compared to both the S&P500 itself and the risk-free asset. They also show a higher standard deviation and a lower sharpe ratio, in both the basic model and the cutoff adjusted one. However, these returns show a lower skew and kurtosis, implying they are closer to normality than the S&P500 returns. In both variations, the model chose to hold a net short position which allowed for a greater than 100% investment in the risk-free asset in 83% of months. Negative cumulative returns would generally imply a poor choice of model, but there are a couple potential explanations. If February and March of 2020 are excluded, the sharpe ratio of the Ex-post returns is 0.14 in both model variations. This finding provides further evidence that drawing a definite conclusion on the cutoff variation vs. the basic model may constitute overfitting. Additionally, while still lower than the S&P500, it could mean that a larger sample would smooth out the Covid volatility. Finally, a deeper analysis of the performance of GARCH and other returns forecasting models is beyond the scope of this paper, and would require a longer series of option data to ensure the results are robust.

This paper leaves several extensions open for future research and implementation to transition the option portfolio optimization model to real world trading. Since option contracts can only be traded in integer quantities as opposed to stocks, making the optimization non-convex by imposing an integer constraint would mean the results could be directly traded on. While increasing the computational complexity, these results could be more effectively tested against named option strategies such as the straddles, spreads, and condors described in Chaput and Ederington (2003), among others. Additionally, using a dynamic rebalancing strategy rather than held to expiration could prove to be more profitable. This formulation does present some additional challenges, however, if a closed form option pricing is not used to derive an analytical expression for the change in portfolio value. In the case of the GARCH model used in this paper, for example, the future value of the portfolio of options would need to be simulated and appropriately discounted. The dynamically rebalanced portfolio does have some advantages over the static hold until expiration strategy, as noted by Zhao and Palomar (2018). Additionally, dynamic rebalancing could occur on any time scale, independent of the expiration date, such as daily or hourly, enabling more trading opportunities. Another extension would be to incorporate further out of the money options, which would increase the choice set in the model, while opening up the issue of low liquidity contracts. To address this, further research could explore creating an effective metric to filter out low liquidity contracts. Lastly, as the skew and kurtosis values of the standardized returns showed only a modest reduction over the log returns in Table 1, the assumption of normality of the standardized returns in the GARCH model in Section 2.2 may not be valid. Other models to simulate the returns of the underlying asset could also be substituted and tested for their predictive accuracy, including variations on the GARCH, time series methods, or neural networks, among others.

References

- Giovanni Barone-Adesi, Robert F Engle, and Loriano Mancini. A garch option pricing model with filtered historical simulation. The review of financial studies, 21(3):1223– 1258, 2008.
- F Black and M Scholes. The pricing of options and corporate liabilities. journal of political economy 81 (may-june). 1973.
- Robert R Bliss and Nikolaos Panigirtzoglou. Option-implied risk aversion estimates. *The journal of finance*, 59(1):407–446, 2004.
- J Scott Chaput and Louis H Ederington. Option spread and combination trading. *The journal of Derivatives*, 10(4):70–88, 2003.
- Bjørn Eraker. The performance of model based option trading strategies. Review of Derivatives Research, 16(1):1–23, 2013.
- José Afonso Faias and Pedro Santa-Clara. Optimal option portfolio strategies: Deepening the puzzle of index option mispricing. *Journal of Financial and Quantitative Analysis*, 52(1):277–303, 2017.
- Markus Gsell. Assessing the impact of algorithmic trading on markets: A simulation approach. Technical report, CFS Working Paper, 2008.
- Lauren A Hannah. Stochastic optimization. International Encyclopedia of the Social & Behavioral Sciences, 2:473–481, 2015.
- Harald Malmgren and Mark Stys. Computerized global trading 24/6. International Economy, Spring, 2011.
- Harry Markowitz. Portfolio selection. The Journal of Finance, 7(1):77–91, March 1952.
- Mogens Graf Plessen and Alberto Bemporad. Parallel investments in multiple call and put

options for the tracking of desired profit profiles. In 2017 American Control Conference (ACC), pages 1091–1096. IEEE, 2017.

- Pedro Santa-Clara and Alessio Saretto. Option strategies: Good deals and margin calls. Journal of Financial Markets, 12(3):391–417, 2009.
- Licheng Zhao and Daniel P Palomar. A markowitz portfolio approach to options trading. *IEEE Transactions on Signal Processing*, 66(16):4223–4238, 2018.

A Options Overview

Options are a type of financial derivative. The two types of basic options, calls and puts, give the purchaser the right but not the obligation to buy or sell, respectively, the underlying security at a future date and specified price. With a European option, the holder of the option can exercise their position at the expiration date if they so choose.¹ The seller, or writer, of the options must fulfill their end of the contract if the option is exercised. As with stocks, buying an option is known as taking a long position, and selling a short position.

The buyer of an option increases their choice set, thereby mitigating risk on the downside for a call, and the upside for a put. The seller takes on the risk that the buyer will exercise their option. Therefore, a buyer should pay the seller a cost for this shifting of risk. The price payed for this exchange of risk is the option price.

The price at which the option holder has the right to buy or sell the underlying security is called the strike price, commonly denoted K. The time in days until maturity of the option, at its expiration date, is denoted T. The price of the option is denoted C and P for calls and puts respectively, as in Faias and Santa-Clara (2017). The price of the underlying asset at expiration is denoted S_T . Option moneyness is defined as $\frac{S_t}{K} - 1$.

Figures 14 and 15 show the profit functions of a long call and a long put respectively. The profit functions of the short side of each of these is obtained by multiplying by -1.

¹An American option, not considered in this paper, gives the option holder the right to exercise at any time. This early exercise ability makes the American option price higher than is European counter part. Closed form option pricing models such as Black Scholes assume European options.





B Detailed ExPost Returns

Table 9:

Detailed Portfolio Optimization Returns

ExPost returns,	and weights for each of the four contracts and the risk free asset ar	е								
shown. All values are in percentages.										

Date	Expiry	ExPost	ATM Call	OTM Call	ATM Put	OTM Put	Risk Free
2019-01-15	2019-02-15	-0.1	-0.1	-0.3	-0.8	-0.8	102.0
2019-02-15	2019-03-15	0.0	-0.7	0.2	0.8	-0.7	100.5
2019-03-18	2019-04-18	-1.8	-1.8	0.3	1.0	-0.8	101.2
2019-04-17	2019-05-17	-0.3	1.8	0.2	1.7	-0.8	97.1
2019-05-21	2019-06-21	2.2	-2.0	0.3	-1.0	-0.3	103.0
2019-06-19	2019-07-19	-2.7	-4.8	0.4	0.0	-0.5	104.8
2019-07-16	2019-08-16	3.5	1.0	0.1	1.6	-1.3	98.5
2019-08-20	2019-09-20	0.0	-3.2	0.8	-2.1	-0.3	104.8
2019-09-18	2019-10-18	2.3	-2.0	0.3	0.3	-0.7	102.1
2019-10-15	2019-11-15	-2.8	-1.8	0.3	-0.0	-0.6	102.2
2019-11-20	2019-12-20	-0.0	0.3	0.3	0.9	-0.9	99.3
2019-12-17	2020-01-17	0.9	1.0	0.0	1.7	-1.2	98.5
2020-01-21	2020-02-21	4.3	-4.4	0.4	0.2	-0.5	104.2
2020-02-20	2020-03-20	-24.2	-23.4	1.9	-1.7	0.1	123.1
2020-03-17	2020-04-17	11.5	-0.2	-1.8	-0.0	-9.6	111.6
2020-04-15	2020-05-15	1.2	-3.9	0.9	-2.6	-0.0	105.7
2020-05-19	2020-06-19	0.5	-1.2	1.1	0.1	-1.0	101.0
2020-06-17	2020-07-17	0.9	-2.3	0.8	-0.0	-1.6	103.1
2020-07-21	2020-08-21	-1.8	-3.5	1.9	1.3	-1.4	101.6
2020-08-18	2020-09-18	-2.8	1.3	0.7	2.7	-1.8	97.2
2020-09-16	2020-10-16	1.4	-1.4	0.3	-0.0	-0.9	102.1
2020-10-20	2020-11-20	2.4	-4.4	-0.0	-3.1	0.2	107.2
2020-11-18	2020-12-18	0.4	-0.9	0.0	0.2	-1.3	102.0
2020-12-15	2021-01-15	-0.5	-1.9	1.1	0.3	-1.4	101.9
2021-01-19	2021-02-19	1.1	-0.9	0.1	0.1	-1.5	102.2
2021-02-19	2021-03-19	1.9	-2.0	1.0	1.4	-2.3	101.9
2021-03-16	2021-04-16	-4.8	-5.3	-0.0	-0.5	-0.6	106.4
2021-04-21	2021-05-21	0.3	-0.3	0.2	0.4	-0.7	100.3
2021-05-18	2021-06-18	1.9	1.0	0.0	-1.0	-0.7	100.7

ExPost returns, and weights for each of the four contracts and the risk free asset are shown for the cutoff variation. All values are in percentages.								
Date	Expiry	ExPost	ATM Call	OTM Call	ATM Put	OTM Put	Risk Free	
2019-01-15	2019-02-15	-0.1	-0.1	-0.3	-0.8	-0.8	102.0	
2019-02-15	2019-03-15	0.0	-0.7	0.2	0.8	-0.7	100.5	
2019-03-18	2019-04-18	-1.8	-1.8	0.3	1.0	-0.8	101.2	
2019-04-17	2019-05-17	-0.3	1.8	0.2	1.7	-0.8	97.1	
2019-05-21	2019-06-21	2.2	-2.0	0.3	-1.0	-0.3	103.0	
2019-06-19	2019-07-19	-2.7	-4.8	0.4	0.0	-0.5	104.8	
2019-07-16	2019-08-16	3.5	1.0	0.1	1.6	-1.3	98.5	
2019-08-20	2019-09-20	0.0	-3.2	0.8	-2.1	-0.3	104.8	
2019-09-18	2019-10-18	2.3	-2.0	0.3	0.3	-0.7	102.1	
2019-10-15	2019-11-15	-2.8	-1.8	0.3	-0.0	-0.6	102.2	
2019-11-20	2019-12-20	-0.0	0.3	0.3	0.9	-0.9	99.3	
2019-12-17	2020-01-17	0.9	1.0	0.0	1.7	-1.2	98.5	
2020-01-21	2020-02-21	4.3	-4.4	0.4	0.2	-0.5	104.2	
2020-02-20	2020-03-20	-28.5	0.0	-1.6	1.8	-1.6	101.4	
2020-03-17	2020-04-17	11.5	-0.2	-1.8	-0.0	-9.6	111.6	
2020-04-15	2020-05-15	1.2	-3.9	0.9	-2.6	-0.0	105.7	
2020-05-19	2020-06-19	0.5	-1.2	1.1	0.1	-1.0	101.0	
2020-06-17	2020-07-17	0.9	-2.3	0.8	-0.0	-1.6	103.1	
2020-07-21	2020-08-21	-1.8	-3.5	1.9	1.3	-1.4	101.6	
2020-08-18	2020-09-18	-2.8	1.3	0.7	2.7	-1.8	97.2	
2020-09-16	2020-10-16	1.4	-1.4	0.3	-0.0	-0.9	102.1	
2020-10-20	2020-11-20	2.4	-4.4	-0.0	-3.1	0.2	107.2	
2020-11-18	2020-12-18	0.4	-0.9	0.0	0.2	-1.3	102.0	
2020-12-15	2021-01-15	-0.5	-1.9	1.1	0.3	-1.4	101.9	
2021-01-19	2021-02-19	1.1	-0.9	0.1	0.1	-1.5	102.2	
2021-02-19	2021-03-19	1.9	-2.0	1.0	1.4	-2.3	101.9	
2021-03-16	2021-04-16	-4.8	-5.3	-0.0	-0.5	-0.6	106.4	
2021-04-21	2021-05-21	0.3	-0.3	0.2	0.4	-0.7	100.3	
2021-05-18	2021-06-18	1.9	1.0	0.0	-1.0	-0.7	100.7	

Table 10:Detailed Portfolio Optimization Returns: Cutoff