



M2 THESIS

Portfolio Optimization with Options

ANDREW BOOMER

September 27, 2021

Table of Contents

1	Abstract	2
2	Introduction	3
2.1	Overview	3
3	Model	6
3.1	Portfolio Optimization Method	6
3.2	Returns	8
3.3	Portfolio Optimization Steps	9
4	Data	12
5	Estimation	17
6	Results	19
6.1	Portfolio Optimization: Initial Model	19
6.2	Portfolio Optimization: Cutoff Adjusted	23
7	Conclusion	26
A	Options Overview	29
B	Detailed ExPost Returns	31

1 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 ad-hoc 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. In line with their work, however, I find that out-of-the money (OTM) calls have an average long position, while at-the-money (ATM) calls, ATM puts, and OTM puts have an average short position.

2 Introduction

2.1 Overview

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 throughout the paper.

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). While under the strict Black-Scholes assumption of no riskless arbitrage, returns from any option portfolio can be replicated with a portfolio of stocks, in the real world, options can achieve excess returns. Additionally, 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 include options in the optimization problem. They use the Black-Scholes assumption of stock returns following

a geometric brownian motion to optimize the set of stock and option 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 aren't 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 dynamically 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. Therefore, another extension for future work would be to reformulate my approach with dynamic rebalancing. This formulation does present some additional challenges 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, however, as noted by Markowitz (1952). Holding until expiration relies on long term return estimation, one month in the case of this paper, which can be unreliable. Additionally, dynamic rebalancing could occur on any time scale, independent of the expiration date, such as daily or hourly, enabling more trading opportunities. Finally, holding until expiration doesn't allow the investor to exit their position if the risk exceeds their tolerance

or the market moves in an unfavorable direction.

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 issues 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 the problem by extension. Convex optimization is more tractable and less computationally expensive than non-convex, so the convex formulation is widely used in

the literature. In a real trading setting, however, option contracts can only be traded in integer quantities, so adding an integer constraint to the set of weights can be an area for extension and further research. This constraint would result in the output of the portfolio optimization model returning actual option combinations used by retail traders and firms, such as straddles, spreads, and condors Chaput and Ederington (2003).

In Section 3 I provide an overview of the portfolio selection method, as well as the GARCH model used to simulate the volatility of the log returns. In Section 4, I describe the data sources, give summary statistics, and present basic performance metrics of the historical data. In Section 5, I present the estimation of the GARCH model. In Section 6, I give the results of the portfolio optimization method. In Section 7, I conclude, review the model and findings, and discuss further research.

3 Model

3.1 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 , we solve:

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

Where F_t as 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 , $r f_t$ 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:

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

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:

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

Transaction costs are modeled through the bid-ask, 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 constrained optimization, 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 Faias and Santa-Clara (2017) cite literature saying transaction costs are partly responsible for pricing anomalies 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 3.3. First, the GARCH volatility model is introduced.

3.2 Returns

In this paper I assume that the log asset returns (y_t) follow a normal distribution with a time varying volatility term. I depart from Faias and Santa-Clara (2017), however, by considering a full GARCH specification, whereas Faias and Santa-Clara (2017) use an ad-hoc approach with relies on estimation of recent realized volatility. Therefore, the full model specification for the log returns of the underlying asset is given by the GARCH (1, 1) with zero conditional mean:

$$y_t = \sigma_t \epsilon_t \quad \epsilon_t \sim iid.N(0, 1)$$

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$

$$\omega > 0 \quad \alpha, \beta \geq 0$$

As explained in further detail in Section 4, 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 $\hat{\sigma}_{t+h} = \sqrt{\hat{\sigma}_{t+h}^2}$

3.3 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 bootstrapped 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:

$$\begin{aligned}
C_{t+1|t}^n(k) &= \max(S_{t+1|t}^n - k, 0) & k &= K_{t,c_1}, \dots, K_{t,c_C} \\
P_{t+1|t}^n(k) &= \max(k - S_{t+1|t}^n, 0) & k &= K_{t,p_1}, \dots, K_{t,p_P} \\
r_{t+1|t,c}^n(k) &= \frac{C_{t+1|t}^n(k)}{C_{t,k}} - 1 & k &= K_{t,c_1}, \dots, K_{t,c_C} \\
r_{t+1|t,p}^n(k) &= \frac{P_{t+1|t}^n(k)}{P_{t,k}} - 1 & k &= K_{t,p_1}, \dots, K_{t,p_P}
\end{aligned}$$

• 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 by concatenating the call/put returns and weights.

$$rp_{t+1|t}^n(\mathbf{W}_t) = rf_t + \langle \mathbf{W}_t, \mathbf{R}_{t+1|t}^n \rangle$$

$$\mathbf{W}_t = [\omega_{t,c_1}, \dots, \omega_{t,c_C}, \omega_{t,p_1}, \dots, \omega_{t,p_P}]$$

$$\mathbf{R}_{t+1|t}^n = [r_{t+1|t,c_1}^n, \dots, r_{t+1|t,c_C}^n, r_{t+1|t,p_1}^n, \dots, r_{t+1|t,p_P}^n] - [rf_t, \dots, rf_t]$$

• Step 5: Present utility function

I use a power utility function which has the property of constant relative risk aversion (CRRA). Faias and Santa-Clara (2017) cite that previous literature estimated γ as 4 for the S&P500 index option data, but I follow their use of $\gamma = 10$. 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.

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

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 $S_{t+1|t}^n$.

$$\begin{aligned} C_{t+1|t}(k) &= \max(S_{t+1} - k, 0) & k &= K_{t,c_1}, \dots, K_{t,c_C} \\ P_{t+1|t}(k) &= \max(k - S_{t+1}, 0) & k &= K_{t,p_1}, \dots, K_{t,p_P} \\ r_{t+1|t,c}(k) &= \frac{C_{t+1|t}(k)}{C_{t,k}} - 1 & k &= K_{t,c_1}, \dots, K_{t,c_C} \\ r_{t+1|t,p}(k) &= \frac{P_{t+1|t}(k)}{P_{t,k}} - 1 & k &= K_{t,p_1}, \dots, K_{t,p_P} \end{aligned}$$

Then vectorizing as before:

$$\begin{aligned} rp_{t+1|t}(\mathbf{W}_t^*) &= rf_t + \langle \mathbf{W}_t^*, \mathbf{R}_{t+1|t} \rangle \\ \mathbf{R}_{t+1|t} &= \begin{bmatrix} r_{t+1|t,c_1}, \dots, r_{t+1|t,c_C}, r_{t+1|t,p_1}, \dots, r_{t+1|t,p_P} \end{bmatrix} - \begin{bmatrix} rf_t, \dots, rf_t \end{bmatrix} \end{aligned}$$

Next periods wealth is then:

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

4 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 $y_t = \ln(S_t) - \ln(S_{t-1})$, 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.

Table 1:
S&P 500 Index Returns Summary Statistics

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).

Statistics	Log Returns y_t			Standardized Returns $\hat{\epsilon}_t$		
	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	19.51	4.65
$\rho_1(z)$	0.24	−0.02	0.22	0.18	−0.01	0.17
$\rho_1(z^2)$	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	0.14
	[0.0]	[0.81]	[0.0]	[0.9]	[0.76]	[0.7]

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.

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.

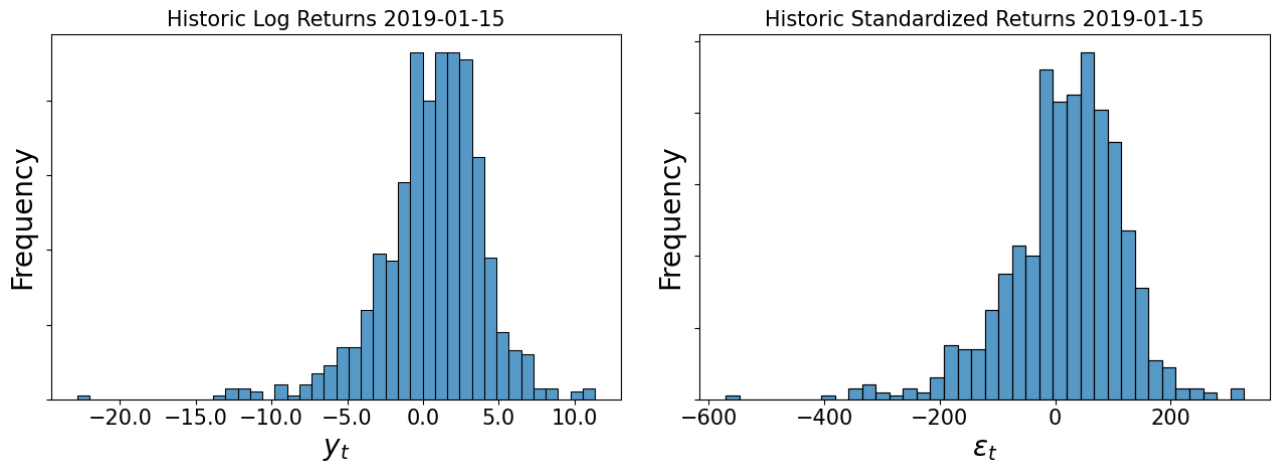
Table 2:
CBOE Variable Descriptions

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

Figure 1:
Historic Log/Standardized Returns

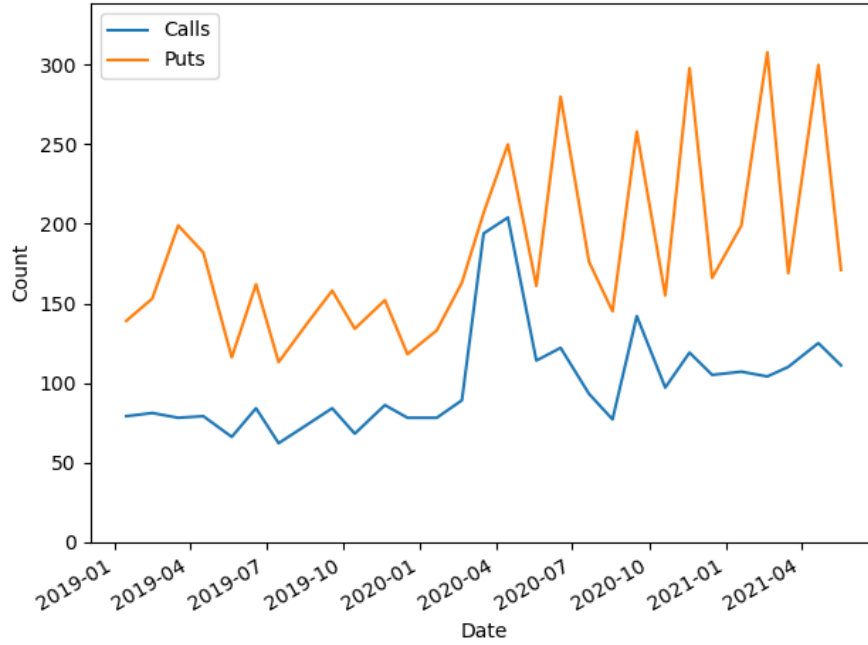
*Histograms of historic log returns (y_t) and standardized returns (ϵ_t) are presented.
Time period spans from 1950 through 2018.*



Observations that meet any of these 3 criteria are removed from the 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

Count of Contracts by date and by option type after CBOE option data was filtered. Time span is from January 2019 to May 2021.



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. 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 3 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 is likely due to the small sample size and the outsized impact of the maximum returns due to the first COVID month (Check!!).

Table 3:
Actual Returns for Options, S&P500, and Risk Free

ATM/OTM Call/Put show individual realized option returns for a hold until expiration strategy. $\frac{1}{N}$ shows return for even weighting on each of the four contracts. S&P500, Risk Free show actual returns for underlying asset and risk free treasury bill respectively. Mean, Std, Min, Max are shown as percentages.

	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	1.8	4.7	-19.1	6.3	-3.14	11.59	0.38
Risk Free	-14.2	31.6	-100.0	44.4	-1.34	1.66	-0.45

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 "DGS1MO"). 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 risk free returns, that as I take as given. Figure 4 shows the time series of the 1-month treasury bill returns. While Figure 4 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 XXX times higher than the risk-free return in the red area, from 2019 through May 2021. As will be discussed in further detail in section 6, this discrepancy may impact a strategy that tends to choose longer positions on the risk free asset.

Figure 3:
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.

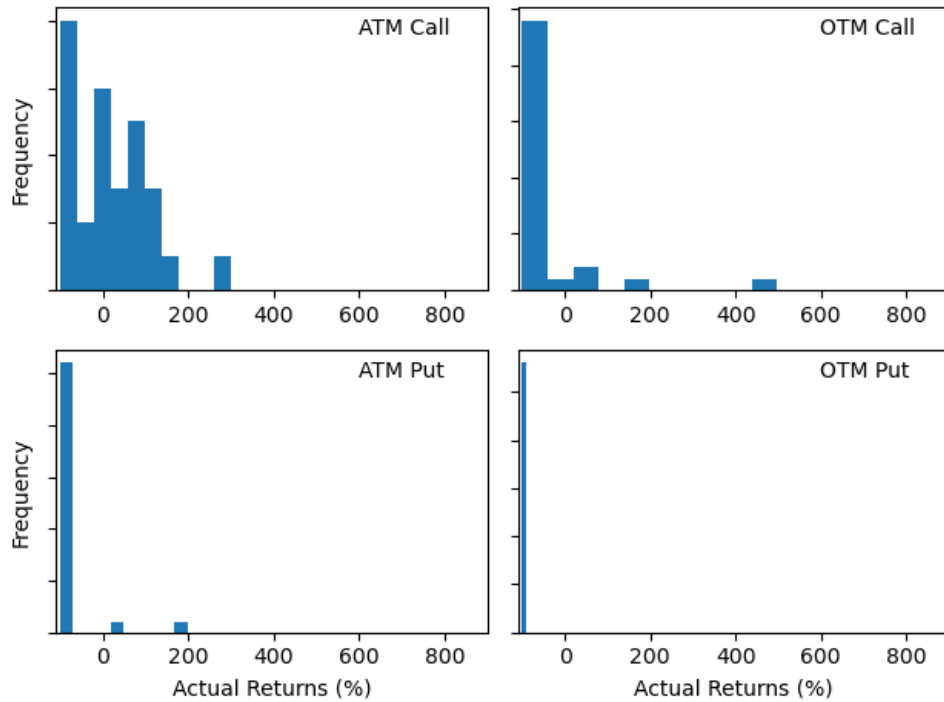
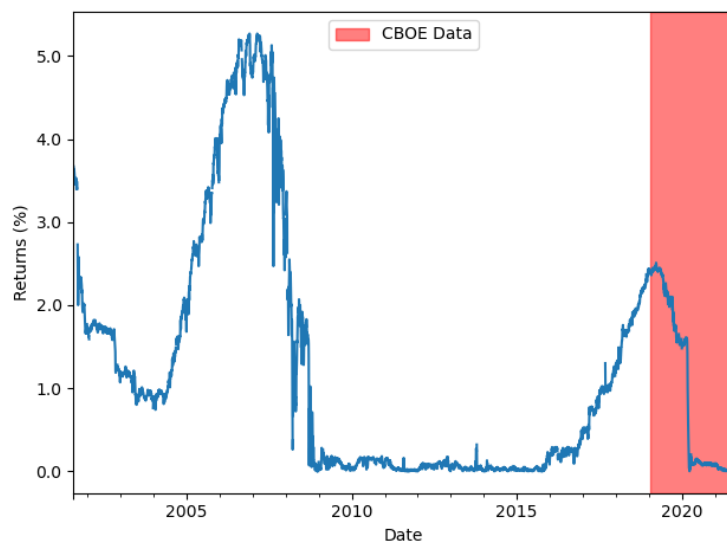


Figure 4:
Risk Free Asset Returns

Time Series of 1-Month Treasury Bill Returns from St. Louis FRED. Time span is from August 2001 to August 2021. Time span of available option data is shown in red.



5 Estimation

The rolling window GARCH(1, 1) model described in section 3.2 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).

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

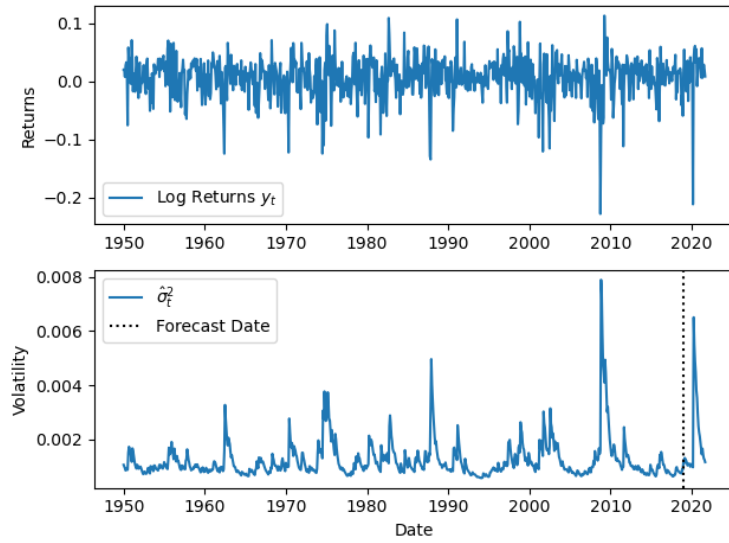
GARCH(1, 1) model estimation summary is shown. The model is estimated on log returns from January 1950 through December 2018.

	coef	std err	t	P> 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

Figure 5 shows a time series plot of the log returns and the conditional volatility using the rolling GARCH procedure described in Section 3.2. The spikes in conditional volatility for the great recession and the Covid pandemic are visible in Figure 5 as well.

Figure 5:
GARCH Time Series Plot

Time series of S&P500 log returns, and estimated conditional volatility from GARCH(1, 1) model are shown. After the forecast date, January 2019, conditional volatility values shown are forecasts. Time spans from January 1950 to May 2021.



6 Results

Expanding on the work done in Faias and Santa-Clara (2017), multiple variations of the portfolio optimization technique laid out in section 3 are presented. The first variation 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 the CRRA parameter γ to 10 does induce some portfolio weight shrinkage, this variation models a scenario where a trader might want to define a cutoff for their single contract risk exposure. The other variation uses all the other option contracts for each date other than the four ATM/OTM calls/puts. This increases the choice set for the portfolio optimization model, which can in theory lead to a more profitable outcome. There are however challenges with including many of the possible option contracts, some of which are described in Faias and Santa-Clara (2017). Introducing many options that are similar in moneyness can lead to redundancy in the optimization, and making trades across many options leads to more exposure to bid-ask spread transaction costs. Additionally, options that are too deep OTM can have low liquidity which exposes an options portfolio to the risk of non-assignment on a portion of the portfolio, or as described in section 2, the possibility of \mathbf{W}_t^* being endogenous to the future options price.

6.1 Portfolio Optimization: Initial Model

Figure 6 shows a histogram of the monthly Ex-post returns. We can see that there is a definite XXX 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 XXX, as shown in Table 5 and in full detail in Table 9 of Appendix B, occurs in March 2020, and the highest monthly return of XXX occurs the next month in April 2020, the lowest return is larger in absolute value than the highest.

Figure 6:
Portfolio Optimization Returns

Histogram of portfolio optimization returns is shown. The time period is January 2019 to May 2021.

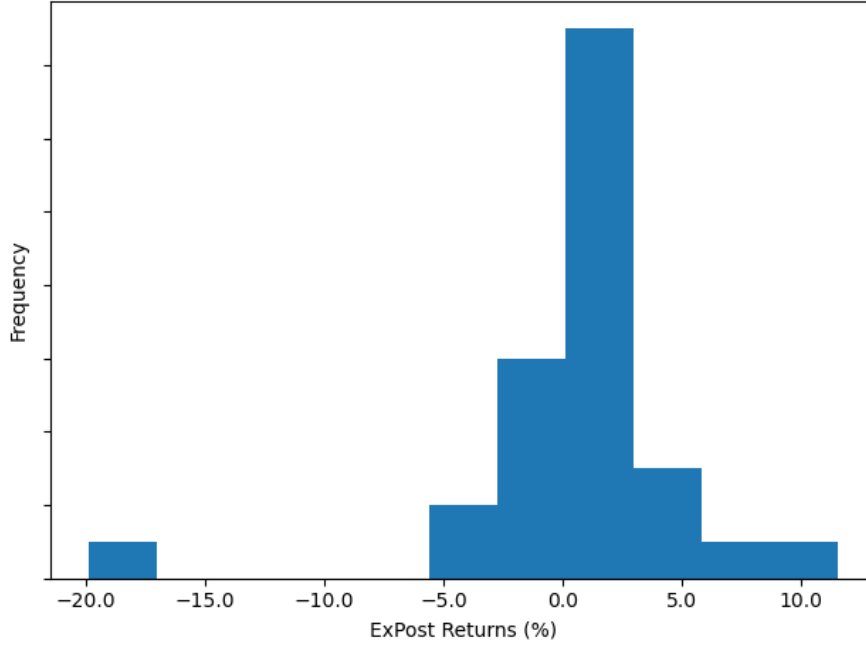


Table 5 compares the summary of the monthly returns of the option portfolio optimization with the returns of the S&P500. The the Ex-post returns have a XXX mean and a XXX standard deviation than the S&P500, resulting in the S&P500 having a XXX Sharpe ratio (XXX vs. XXX) in time period of the data. Additionally, the Ex-post returns have a XXX skew and a XXX kurtosis, implying that they are XXX to a normal distribution than the S&P500 returns.

Table 5:
Portfolio Optimization vs. S&P500 Returns

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 percentages. The Sharpe Ratio (SR) is monthly.

	Mean	Std	Min	Max	Skew	Kurtosis	SR
S&P 500	1.8	4.7	-19.1	6.3	-3.14	11.59	0.38
ExPost	0.7	5.0	-19.9	11.5	-2.22	9.15	0.15

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 4, the model is long OTM calls on average, while short on

average for the other three contract types. 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 7 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 XXX% to XXX% in March 2020, due to a large short position in the ATM call (XXX%). On average, the portfolio optimization model is a net XXX of options, and is in a XXX position XXX% of the months. In line with the findings in Faias and Santa-Clara (2017), the model overwhelmingly shorts OTM puts, XXX% of the months, with an average short position of XXX%.

Table 6:
Portfolio Optimization Weights

Mean, min, and max are presented for the optimized portfolio weights. The time period is January 2019 to May 2021. All values are percentages.

	ATM Call	ATM Put	OTM Call	OTM Put
Mean	-2.4	-0.0	0.5	-1.1
Minimum	-25.4	-3.0	-1.9	-9.5
Maximum	1.7	2.9	2.1	0.3

Figure 7:
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 asset. The time period is January 2019 to May 2021.

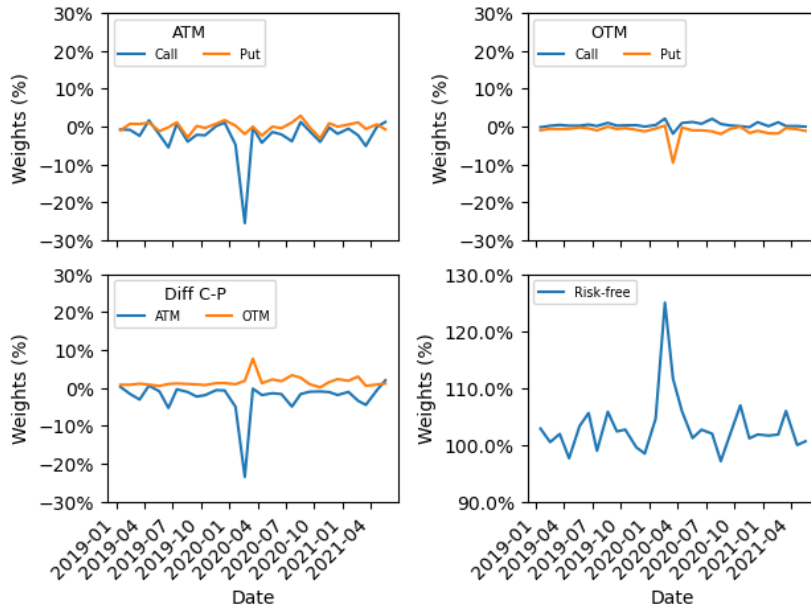
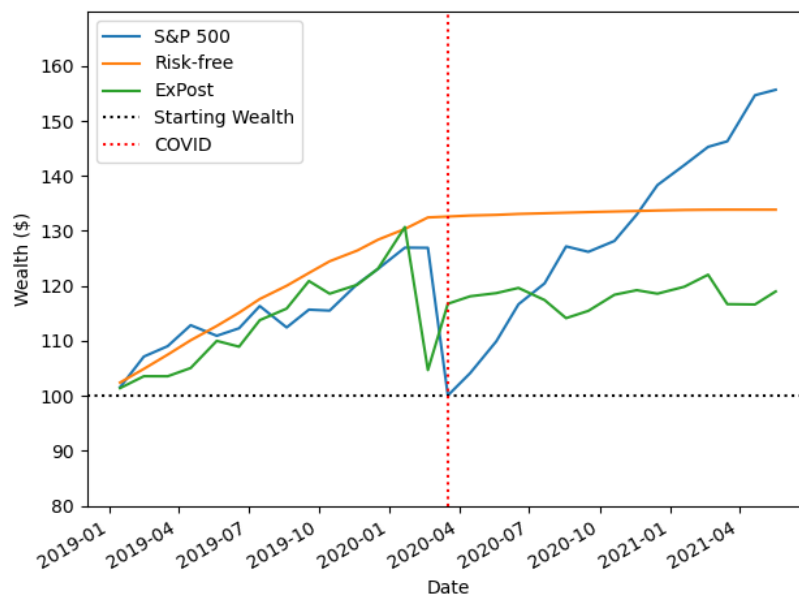


Figure 8 shows the cumulative returns from January 2019 to May 2021 for the portfolio

optimization model, the S&P500, and the risk-free asset. The initial wealth and March 2020 are marked with dotted lines. Prior to Covid, the cumulative returns of all three track relatively closely, however, each respond to the Covid shock differently. The risk-free asset levels out, in accordance with Figure 4, while the Ex-post returns and the S&P500 both experience steep losses, with the S&P500 experiencing a XXX loss. However, the S&P500 bounces back more quickly, and has the highest ending cumulative return of the three.

Figure 8:
Cumulative Returns

Cumulative returns for portfolio optimization, S&P500, and risk-free asset are shown. The time period is January 2019 to May 2021, and starting wealth is set to \$100.



6.2 Portfolio Optimization: Cutoff Adjusted

Figure 9:
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.

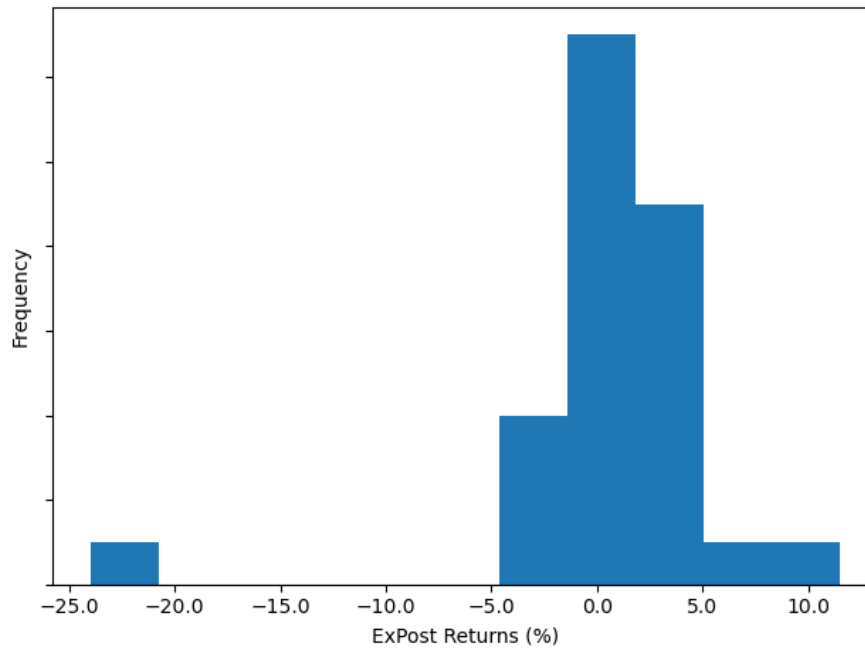


Table 7:
Portfolio Optimization vs. S&P500 Returns: Cutoff Adjusted

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 percentages. The Sharpe Ratio (SR) is monthly. Simulated returns for contracts with weights greater than 10% in absolute value are forced to zero, and the optimization is rerun.

	Mean	Std	Min	Max	Skew	Kurtosis	SR
S&P 500	1.8	4.7	-19.1	6.3	-3.14	11.59	0.38
ExPost	0.6	5.6	-24.0	11.5	-2.76	11.29	0.11

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	-1.7	0.1	0.4	-1.2
Minimum	-5.6	-3.1	-1.5	-9.9
Maximum	1.7	2.8	1.9	0.0

Figure 10:
Cumulative Returns: Cutoff Adjusted

Cumulative returns for portfolio optimization, S&P500, and risk-free asset are shown. The time period is January 2019 to May 2021, and starting wealth is set to \$100. Simulated returns for contracts with weights greater than 10% in absolute value are forced to zero, and the optimization is rerun.

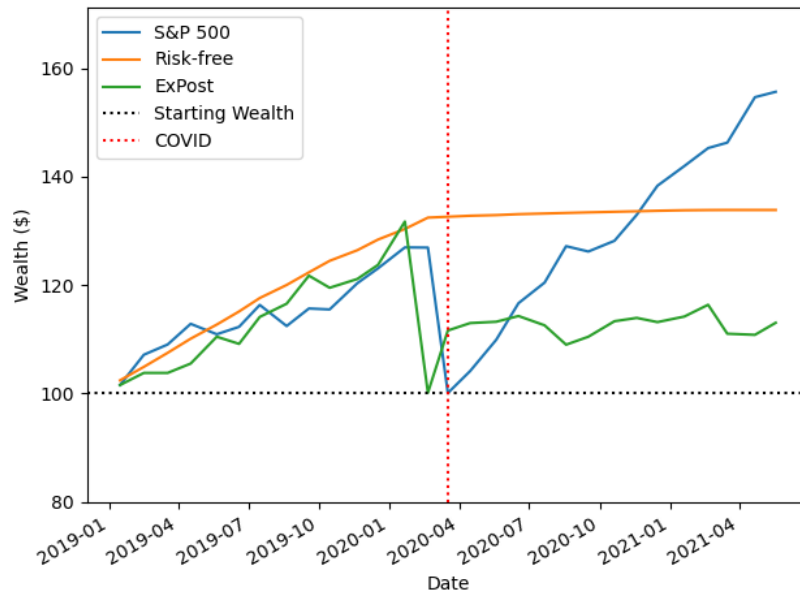
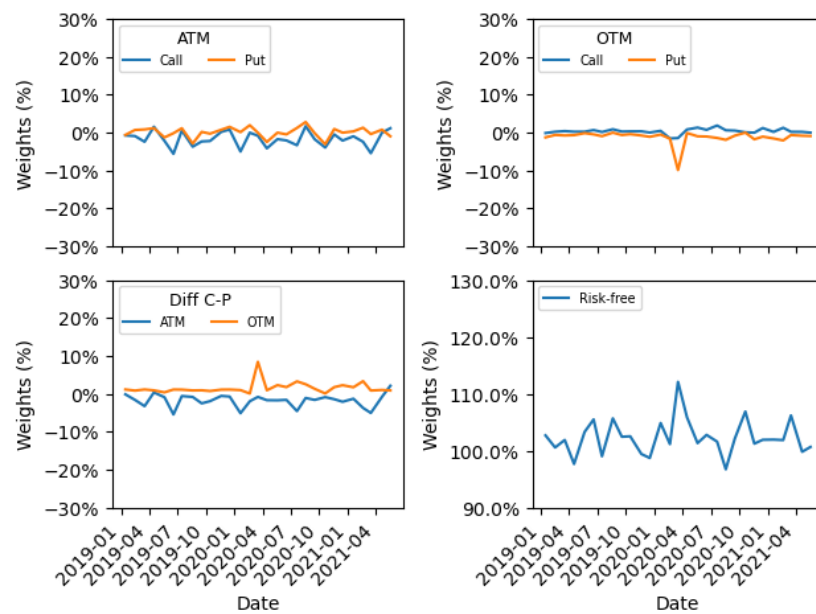


Figure 11:
**Portfolio Optimization Weights by Contract: Cutoff
Adjusted**

Optimized weights from portfolio optimization are shown by option type, as well as for the call-put difference and risk-free asset. 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.



7 Conclusion

The results portfolio optimization model presented in this paper show ex-post returns which are on average higher compared to the S&P500 itself. However, these returns show a lower sharpe ratio, higher kurtosis, and a higher skew in absolute value. Given the shorter time period of the available options data, and the anomalous time period during the COVID pandemic, replicating the optimization on the same options data as in Faias and Santa-Clara (2017) would provide a more accurate comparison. Of particular interest, however, was the ability of the model to accurately predict the profitability of heavily shorting ATM puts during the beginning of the COVID pandemic.

This paper leaves several extensions open for future research and implementation to transition the option portfolio optimization model to real world trading. Imposing an integer constraint on the weights would mean the results could be directly traded on. Implementing this model as a dynamically rebalanced portfolio, rather than held to expiration could prove to be more profitable. Incorporating further out of the money contracts would make this optimization more realistic, while however, opening up the issue of low liquidity contracts. Other models to simulate the returns of the underlying asset can also be substituted, including variations on the GARCH, as well as neural networks.

References

- 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.
- 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.
- Licheng Zhao and Daniel P Palomar. A markowitz portfolio approach to options trading. *IEEE Transactions on Signal Processing*, 66(16):4223–4238, 2018.
- Lauren A Hannah. Stochastic optimization. *International Encyclopedia of the Social & Behavioral Sciences*, 2:473–481, 2015.
- Markus Gsell. Assessing the impact of algorithmic trading on markets: A simulation approach. Technical report, CFS Working Paper, 2008.
- J Scott Chaput and Louis H Ederington. Option spread and combination trading. *The journal of Derivatives*, 10(4):70–88, 2003.
- Giovanni Barone-Adesi, Robert F Engle, and Lorian Mancini. A garch option pricing model with filtered historical simulation. *The review of financial studies*, 21(3):1223–1258, 2008.
- Bjørn Eraker. The performance of model based option trading strategies. *Review of Derivatives Research*, 16(1):1–23, 2013.
- 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.

F Black and M Scholes. The pricing of options and corporate liabilities. journal of political economy 81 (may-june). 1973.

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 paid 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 12 and 13 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 counterpart. Closed form option pricing models such as Black Scholes assume European options.

The Profit is

$$\max(X_t - K_i, 0) - C_i = (X_t - K_i)^+ - C_i$$

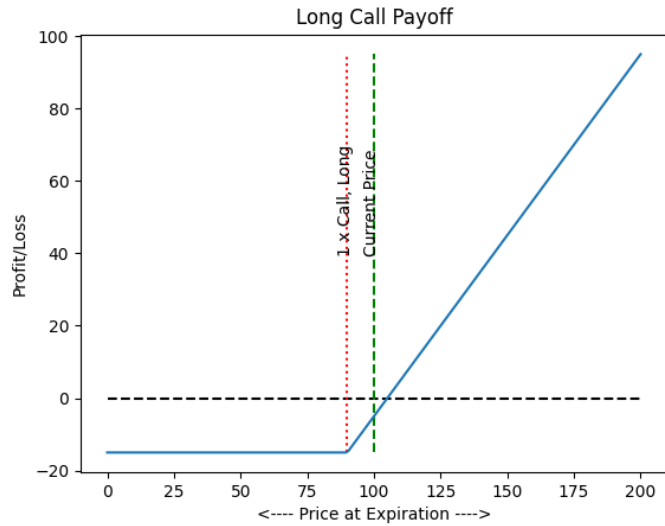


Figure 12

The Profit is

$$\max(K_i - X_t, 0) - P_i = (K_i - X_t)^+ - P_i$$

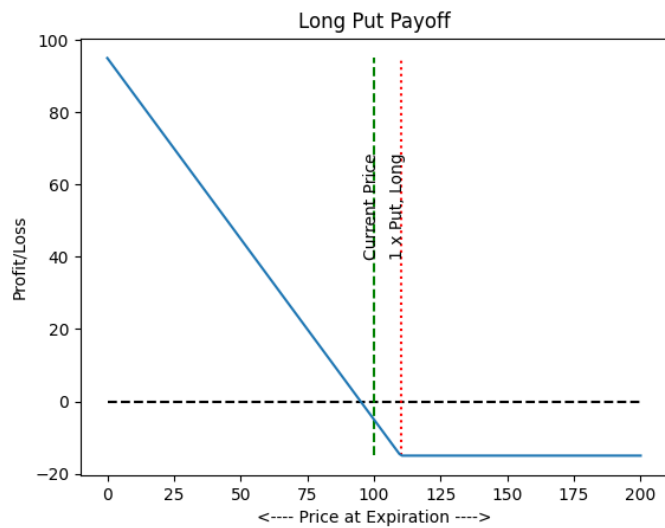


Figure 13

B Detailed ExPost Returns

Table 9:
Detailed Portfolio Optimization Returns

Date	Expiry	ExPost	ATM Call	OTM Call	ATM Put	OTM Put	Risk Free
2019-01-15	2019-02-15	1.7	-0.7	-0.1	-0.7	-1.2	102.7
2019-02-15	2019-03-15	2.3	-0.9	0.2	0.7	-0.6	100.6
2019-03-18	2019-04-18	0.2	-2.2	0.4	0.8	-0.7	101.8
2019-04-17	2019-05-17	1.7	1.7	0.2	1.4	-0.8	97.5
2019-05-21	2019-06-21	4.6	-2.2	0.3	-1.0	-0.3	103.2
2019-06-19	2019-07-19	-1.4	-5.9	0.8	-0.3	-0.4	105.9
2019-07-16	2019-08-16	5.4	0.6	0.2	1.4	-1.3	99.0
2019-08-20	2019-09-20	1.9	-4.0	1.1	-2.9	-0.1	105.8
2019-09-18	2019-10-18	4.6	-2.3	0.3	0.0	-0.5	102.6
2019-10-15	2019-11-15	-1.6	-2.2	0.3	-0.4	-0.4	102.6
2019-11-20	2019-12-20	1.5	0.2	0.3	0.8	-0.9	99.6
2019-12-17	2020-01-17	2.1	0.9	0.1	1.7	-1.2	98.5
2020-01-21	2020-02-21	6.5	-5.0	0.5	0.2	-0.5	104.9
2020-02-20	2020-03-20	-23.8	-25.5	2.1	-1.9	0.2	125.1
2020-03-17	2020-04-17	11.5	-0.7	-1.6	-0.0	-9.7	112.0
2020-04-15	2020-05-15	1.1	-4.5	1.1	-2.6	-0.0	106.0
2020-05-19	2020-06-19	0.5	-1.3	1.2	0.0	-0.9	101.0
2020-06-17	2020-07-17	0.7	-1.8	0.7	-0.3	-1.0	102.3
2020-07-21	2020-08-21	-1.9	-4.0	2.2	1.1	-1.3	101.9
2020-08-18	2020-09-18	-3.0	1.4	0.7	3.0	-2.0	96.9
2020-09-16	2020-10-16	1.3	-1.6	0.4	-0.0	-0.9	102.0
2020-10-20	2020-11-20	2.2	-4.6	0.3	-3.0	0.0	107.2
2020-11-18	2020-12-18	0.7	-0.5	-0.1	0.2	-1.2	101.5
2020-12-15	2021-01-15	-0.5	-1.9	1.1	-0.1	-1.0	101.8
2021-01-19	2021-02-19	1.0	-0.7	0.2	0.1	-1.3	101.7
2021-02-19	2021-03-19	1.6	-2.0	1.1	1.0	-1.6	101.6
2021-03-16	2021-04-16	-4.4	-5.2	0.2	-0.5	-0.5	105.9
2021-04-21	2021-05-21	-0.1	0.0	0.2	0.6	-0.6	99.9
2021-05-18	2021-06-18	2.3	1.3	0.0	-1.0	-1.0	100.8

Table 10:
Detailed Portfolio Optimization Returns: Cutoff

Date	Expiry	ExPost	ATM Call	OTM Call	ATM Put	OTM Put	Risk Free
2019-01-15	2019-02-15	1.7	-0.7	-0.1	-0.7	-1.2	102.7
2019-02-15	2019-03-15	2.3	-0.9	0.2	0.7	-0.6	100.6
2019-03-18	2019-04-18	0.2	-2.2	0.4	0.8	-0.7	101.8
2019-04-17	2019-05-17	1.7	1.7	0.2	1.4	-0.8	97.5
2019-05-21	2019-06-21	4.6	-2.2	0.3	-1.0	-0.3	103.2
2019-06-19	2019-07-19	-1.4	-5.9	0.8	-0.3	-0.4	105.9
2019-07-16	2019-08-16	5.4	0.6	0.2	1.4	-1.3	99.0
2019-08-20	2019-09-20	1.9	-4.0	1.1	-2.9	-0.1	105.8
2019-09-18	2019-10-18	4.6	-2.3	0.3	0.0	-0.5	102.6
2019-10-15	2019-11-15	-1.6	-2.2	0.3	-0.4	-0.4	102.6
2019-11-20	2019-12-20	1.5	0.2	0.3	0.8	-0.9	99.6
2019-12-17	2020-01-17	2.1	0.9	0.1	1.7	-1.2	98.5
2020-01-21	2020-02-21	6.5	-5.0	0.5	0.2	-0.5	104.9
2020-02-20	2020-03-20	-28.6	0.0	-1.7	1.9	-1.7	101.4
2020-03-17	2020-04-17	11.5	-0.7	-1.6	-0.0	-9.7	112.0
2020-04-15	2020-05-15	1.1	-4.5	1.1	-2.6	-0.0	106.0
2020-05-19	2020-06-19	0.5	-1.3	1.2	0.0	-0.9	101.0
2020-06-17	2020-07-17	0.7	-1.8	0.7	-0.3	-1.0	102.3
2020-07-21	2020-08-21	-1.9	-4.0	2.2	1.1	-1.3	101.9
2020-08-18	2020-09-18	-3.0	1.4	0.7	3.0	-2.0	96.9
2020-09-16	2020-10-16	1.3	-1.6	0.4	-0.0	-0.9	102.0
2020-10-20	2020-11-20	2.2	-4.6	0.3	-3.0	0.0	107.2
2020-11-18	2020-12-18	0.7	-0.5	-0.1	0.2	-1.2	101.5
2020-12-15	2021-01-15	-0.5	-1.9	1.1	-0.1	-1.0	101.8
2021-01-19	2021-02-19	1.0	-0.7	0.2	0.1	-1.3	101.7
2021-02-19	2021-03-19	1.6	-2.0	1.1	1.0	-1.6	101.6
2021-03-16	2021-04-16	-4.4	-5.2	0.2	-0.5	-0.5	105.9
2021-04-21	2021-05-21	-0.1	0.0	0.2	0.6	-0.6	99.9
2021-05-18	2021-06-18	2.3	1.3	0.0	-1.0	-1.0	100.8