Mutual Information for Optimal Asset Allocation

K. C. CHANG
ZHI TIAN

Dynamic asset allocation in financial investment with an optimal equity growth principle based on mutual information in communication theory is considered. Specifically, the asset allocation formula using Kelly's criteria derived from channel capacity of a binary symmetric channel is developed. The goal is to determine the optimal fraction of equity to be invested between a risk-free asset and a risky asset in a repeated trading activity. The analytical operating curve to predict trading performance is provided. An extension for dynamic multi-asset allocation is also presented. An out-of-sample simulation based on historical market data demonstrates the effectiveness of the methodology.

Manuscript received December 21, 2016; revised June 8, 2017; released for publication December 8, 2017.

Refereeing of this contribution was handled by Ramona Georgescu.

Authors' addresses: K. C. Chang, Dept. of Systems Engineering and Operations Research, George Mason University, Fairfax, VA, U.S.A. (E-mail: kchang@gmu.edu). Z. Tian, Dept. of Electrical and Computer Engineering, George Mason University, Fairfax, VA, U.S.A. (E-mail: ztain1@gmu.edu).

1557-6418/18/$17.00 © 2018 JAIF

1. INTRODUCTION

In modern financial investment world, dynamic asset allocation allows frequent rebalancing of portfolio over time in order to achieve certain objectives. Investors have to sift through a large amount of data in order to analyze the market behavior, predict future market directions, and make sound trading decisions. Given the complexity of the markets and the high stake of trading decisions, financial engineering and risk analysis have emerged as an important research field [1–2].

In particular, one of the critical questions is how to allocate the capital optimally among various correlated risky assets in order to achieve the highest overall return under a defined risk level. This type of portfolio management has become one of the most important elements in practical investment management. The main goal of asset allocation is to develop a long-term risk and return expectation curve for the portfolio and to establish an operating point for each individual investor to balance between the expected return and risk based on his or her own preferences. Traditionally, the process of asset allocation is to identify fundamentally different core asset classes (stocks, bonds, real estate, commodity, etc.) and decide what portion of the capital to invest in each class in order to compose an overall balanced portfolio.

To model and analyze financial data, many mathematical and statistical methods have been applied for quantitative analysis, such as time series analysis, regression analysis, machine learning methods, and Monte Carlo simulations [1–2]. In the financial markets, there are two main traditional approaches for market analysis and stock selection. Fundamental analysis looks into economic factors such as financial statements and market competitiveness to make subjective judgments on the qualitative relationship between equity values and expected market returns, whereas technical analysis uses quantitative historical data of a security such as trading patterns and volume to predict its future price movement [3–8].

Alternative to these traditional approaches, modern quantitative analysis applies complex mathematical models to analyze portfolio risk and develop algorithm trading and arbitraging strategies [2]. This paper adopts the concept of the classical modern portfolio theory (MPT) [9–11], which provides a foundation for explicit risk-reward trade-off analysis. In MPT, a portfolio is consisted of a set of correlated assets each with its own expected return (defined as annual average percentage return) and risk (defined as equity return standard deviation on an annual basis). The goal of this "mean-variance" (MV) approach is to allocate equity among the assets optimally so that the expected portfolio return can be maximized given a defined risk level or vice versa (the overall portfolio risk can be minimized given a desired return level). MPT develops a set of optimal asset allocation policies by optimizing the allocation among available assets subject to risk constraints. These policies form an "efficient frontier" which allows the decision maker to select his/her own operating point on the curve to trade off between risk and return. When an asset is selected with a certain allocation, it implies that specific trading actions for the underlined asset would be taken.

Traditionally MPT-based performance prediction is made based on some idealized dynamic model for the volatility of the underlying asset [11–12], which does not work well in real world. Another approach developed originally by Kelly [13] shows that the optimal long run asset allocation strategy can be obtained by maximizing the expected logarithm of the portfolio value over each time step. This strategy has been proven in different ways [14–15] and has been successfully applied in financial markets [16–17]. An explicit connection between Kelly's criterion and the information theory has also been discussed in [18].

In real world financial market, characteristics of asset returns change rapidly over time such that an investor needs to develop a dynamic asset allocation/rebalancing strategy adaptive to the market environment [5, 19]. In portfolio rebalancing, when an asset is selected with a certain allocation, it implies that specific trading actions for the underlined asset would be taken. Typically, equity-trading strategies are simple buy or sell actions, which are often used for short-term trading.

To account for the complex dynamics of modern market behavior, this paper presents an approach for performance prediction and evaluation by incorporating historical market data in simulated trading. Specifically, the equity index futures and options (S&P 500) data from 1990 to 2016 are used to test the allocation strategy and trading decisions. In addition, instead of defining risk using volatility as in MPT, a different risk metric based on probability of ruin or "draw down" is computed to derive the corresponding efficient frontier.

In this paper, we advocate the development of optimal asset allocation strategies using Kelly's formula that was derived based on mutual information of a binary symmetric channel in communication theory. For clear exposition, we start with a simple trading scenario where the entire equity is allocated between two assets: the risk-free asset (money market or fixed income) and a risky equity asset (S&P index futures and options). In fact, the trading choices examined here are intentionally simplified so that we can clearly illustrate the analytical performance prediction with tradeoffs between risk and return.

The goal is to identify the optimal fraction of the equity to be allocated for trading in order to achieve the highest long-term return given a risk constraint. The results based on extended simulated trading with index options are compared with the analytical performance prediction. In the trading process, in order to obtain a more realistic options price, we develop an analytical model based on "implied volatility"[1] and adjust the option price accordingly. The results have been validated against the historical market data and proved to be reasonably accurate.

The preliminary version of this paper was presented in the 19th International Conference on Information Fusion [20]. We have thoroughly re-organized and revised the original paper with additional contributions. In particular, we develop a dynamic asset allocation strategy for portfolio with multiple correlated assets. Similar to Kelly's approach for single asset, the goal is to maximize the long-run portfolio growth rate over many investment cycles. In addition, we conduct extensive out-of-sample simulations and show that the resulting strategy outperforms the traditional MV approach in expected return but with higher volatility. This strategy can be considered as an alternative approach for the investor to trade off between risk and reward.

The paper is organized as follows. Section 2 describes the optimal asset allocation methodologies and the implications on risk and return. Section 3 presents the selected trading strategies and the options pricing model. Section 4 summarizes the simulation results and the performance analysis for single asset allocation. Section 5 presents Kelly's approach as a multi-asset dynamic allocation strategy and compares its performance with a naïve strategy and MV tangency portfolio. Conclusions and future research are presented in Section 6.

2. OPTIMAL ASSET ALLOCATION

The goal of constructing an optimal portfolio is to maximize the investor's return with a given risk level. Consider a portfolio consisting of multiple assets. The log return of each asset in the portfolio over an investment period is defined as $\eta_i(k) = \log[A_{i,k}/A_{i,k-1}]$ where $A_{i,k}$ is the value of the asset i at time k. Assume that the single period log returns are independent and normally distributed. Then the return of the asset,

$$R_i(k) = (A_{i,k} - A_{i,k-1})/A_{i,k-1} = e^{\eta_i(k)} - 1 \qquad (1)$$

is log-normally distributed.

2.1. Mean-Variance Approach

In modern portfolio theory (MPT) [21–22], a portfolio consists of a set of correlated assets each with its own expected return (defined as annualized percentage return) and risk (defined as equity standard deviation on an annual basis). For example, let $\mathbf{R} = [R_1 \cdots R_N]^T$ be the return of N correlated assets, $E(R) = \boldsymbol{\mu} = [\mu_1 \cdots \mu_N]^T$ be their expected return over an investment period, and Σ be the covariance matrix of $\mathbf{R}$. Denote $\boldsymbol{\omega} = [\omega_1 \cdots \omega_N]^T$ as the asset weights in the portfolio such that $\sum_{i=1}^{N} \omega_i = \mathbf{1}^T \boldsymbol{\omega} = 1$. The expected portfolio return is therefore $\sum_{i=1}^{N} \omega_i \mu_i = \boldsymbol{\omega}^T \boldsymbol{\mu}$ and the corresponding variance is $\sigma_P^2 = \boldsymbol{\omega}^T \Sigma \boldsymbol{\omega}$.

[1]The volatility of the option implied by the market price with a theoretical option value model.

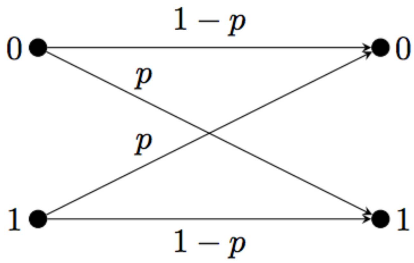


Fig. 1. A Binary Symmetric Channel

Typical investment goal is to allocate capitals among the assets optimally so that the expected portfolio return can be maximized given a defined risk level; or alternatively, the overall portfolio risk can be minimized given a desired return level. For example, if there is no risk-free asset, a typical asset allocation problem can be formulated as the following constrained optimization problem:

$$\text{Min}_{\omega}\, \boldsymbol{\omega}^T \Sigma \boldsymbol{\omega} \text{ subject to } \boldsymbol{\omega}^T \mathbf{1} = 1 \text{ and } \boldsymbol{\omega}^T \boldsymbol{\mu} = \mu_P \quad (2)$$

With this "mean-variance" (MV) approach, an efficient frontier can be constructed where for each return level a portfolio can be derived with minimum risk (variance).

2.2. Binary Symmetric Channel and the Kelly Criteria

In communication theory, a binary symmetric channel (BSC) is defined in Figure 1, where the binary input experiences a cross-over probability (probability of error) p to yield the binary output. In BSC, the channel capacity is defined by

$$C_{\text{BSC}} = 1 - H(p) = 1 + p\log(p) + (1-p)\log(1-p) \quad (3)$$

which is equivalent to the *maximum mutual information*.

The Kelly criterion was originally developed [13] based on the channel capacity concept in communication theory. Start with a single asset portfolio, the specific question addressed was how to allocate the asset optimally for an investment/betting opportunity in order to maximize the expected long-term equity growth rate. The only requirement is that the investment opportunity needs to have a positive expected return (i.e., with a winning edge).

Consider a specific investment (bet). Let p represent its winning probability, b represent the expected return per unit bet for a winning trade, f represent the fraction of the equity allocated to the investment (the remaining $1-f$ sits on the side line). The number of winning and losing trades over n bets is denoted by W and L respectively, with $W + L = n$. Apparently, W and L approach pn and $(1-p)n$ respectively when n is large. Let X_0 and X_n denote the initial and the final amount of the equity after n bets, where X_n/X_0 is called the terminal wealth ratio (TWR). Then the expected log growth rate of the equity per trade can be written as [20],

$$\begin{aligned} g(f) &= E\left\{\log\left[\frac{X_n}{X_0}\right]^{1/n}\right\} \\ &= E\left\{\frac{W}{n}\log(1+bf) + \frac{L}{n}\log(1-f)\right\} \end{aligned} \quad (4)$$

It can be easily shown that the optimal f that maximize $g(f)$ is,

$$f^* = \frac{bp - (1-p)}{b} \quad (5)$$

which is called the Kelly's formula. Specifically, when $b = 1$, $f^* = 2p - 1$, Equation (4) converts to (3), the BSC channel capacity or the *maximum mutual information* [14–15]. In other words, with the optimal fraction allocation based on the Kelly's formula (5), the expected log growth rate of the equity per trade with a winning probability p converges to the maximum mutual information of a binary symmetric channel.

2.3. Return and Risk Trade-off

Equation (4) shows that the log equity of a portfolio is expected to grow at a rate of $g(f)$ on a per-trade basis. For example, if $b = 1$, $p = 0.6$, then the optimal Kelly's fraction is $f^* = 0.2$ and the highest expected gross return per unit bet per trade is $bp - (1-p) = 0.2$. The corresponding expected growth rate of the equity is $R = e^{g(f^*)} - 1 = 0.02$ on a per-trade basis. In other words, with the optimal Kelly criterion, the equity is expected to grow on an average of 2% after each trade if 20% of the equity is allocated for each trade where the amount of potential gain or loss are the same for each trade and the winning probability of each trade is 0.6.

While the Kelly's formula provides an optimal allocation for each trade to maximize the long-term TWR, the potential risk could also be high. Traditionally, trading risk is assessed by the volatility (STD) of the equity return [9]. However, it has been recognized that a more appropriate measure of risk for an investment is draw down (DD) or probability of ruin. Draw down is defined as the percentage of the equity loss from a peak to a subsequent bottom within an investment cycle and the probability of ruin is defined as the probability of losing a certain percentage of the equity at the end of an investment cycle. While return and risk go against each other, it is obviously desirable for an investment strategy to achieve a high return with a limited draw down or low probability of ruin.

It has been shown that with Kelly's investment strategy, the probability of ruin at the terminal stage (having a draw down of $d = 1 - \bar{d}$) after an extensive trading period can be approximated by [15]

$$P\{(X_n/X_0) \le \bar{d}\} = \bar{d}^{(2/c)-1} \quad (6)$$

where the allocation for each trade is $f = cf^*$ with $0 < c < 2$. For example, with $f^* = 0.5$, $c = 1$, and $\bar{d} = 0.33$,

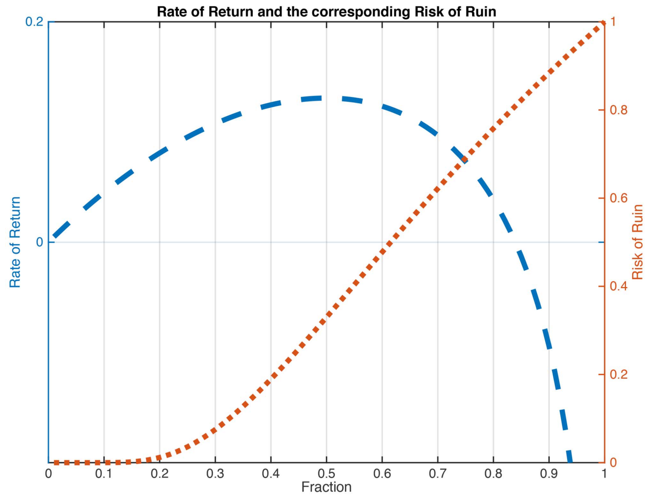


Fig. 2. Risk and Return Trade-off as a function of Investment Fraction

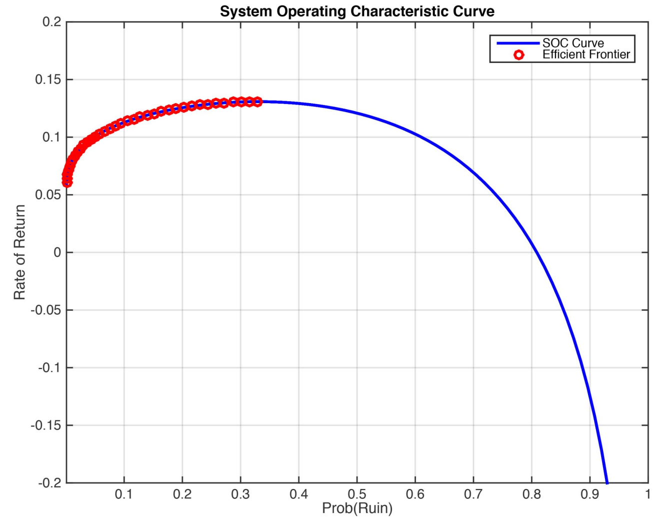


Fig. 3. System Operating Curve (SOC) for Return vs. Risk

$P\{(X_n/X_0) \leq 0.33\} = 0.33$ when $n \gg 1$. Namely, there is a 33% chance that the terminal equity after a large number of trades is less than 33% (a draw down of more than 67%) of the initial equity if 50% (optimal Kelly fraction) of the equity was allocated for each trade. On the other hand, if $c = 0.5$, $f = 0.5f^* = 0.25$ (half of Kelly), then the probability of ruin reduces to $P\{(X_n/X_0) \leq 0.33\} = 0.33^3 \approx 0.036$.

For example, with $p = 0.75$, $b = 1$, and $f^* = 0.5$, Figure 2 compares the normalized rate of return to the risk of ruin over the range of fraction f from 0 to 1 for equity allocation in each trade [20]. As can be seen from the figure, when the allocation follows the Kelly's suggestion (namely, $f = 0.5$), the rate of return is at maximum (0.13) while the probability of ruin (with $\bar{d} = 0.33$), is about 0.33. When half of Kelly (0.25) is applied, the rate of return lowers to 0.095 while the risk of ruin reduces to under 0.04. This also suggests that a systematic approach can be developed for an investment strategy where a system operating curve (SOC) can be established to predict risk and reward performance at different operating points and ultimately allow an investor to choose a specific point to fit his/her own risk preference.

To illustrate, Figure 3 shows the corresponding SOC curve derived from Figure 2 [20]. Each point on the curve represents an operating point with a specific fraction of equity being allocated for each trade, where the peak of the curve corresponds to the suggested optimal Kelly's fraction (50%). The left portion of the curve represents the operating points where higher return can be achieved by taking a higher risk. They can be considered as the "investment efficient frontier" where investors could pick and choose an operating point based on their own preference. It can be seen from the figure that beyond the efficient frontier, taking higher risk will negatively reduce the expected rate of return. This is due to the "over aggressiveness" by investing more than the optimal Kelly's fraction on each trade and is clearly not desirable.

3. TRADING STRATEGIES WITH SINGLE ASSET ALLOCATION

Having explained the basic principle of Kelly's formula, this section demonstrates how it can be applied for optimal allocation with a single asset in practice where a number of statistical characteristics of the trading asset need to be acquired first. Specifically, the winning probability, the average gain of winning trades, and the average loss of the losing trades need to be estimated.

3.1. Trading S&P Futures and Options

S&P futures and their options are selected as the asset for the portfolio. They are traded in many financial markets, including the Chicago Mercantile Exchange (CME) [23] and the CME electronic GLOBEX platform [24]. They are one of the most liquid equity index products traded in the world. The two most commonly traded options are the plain vanilla put and call options.[2] Both S&P futures and the corresponding options are extremely liquid and popular.

One could long or short the futures contracts or the option contracts depending on the goals of his/her trading strategies. By writing (selling) the put options when the market is expected to go higher would result in the options expiring worthlessly and therefore the seller could keep the collected premium. Similarly, the seller could keep the premium collected by writing the call options if the market does not move up beyond the strike price. However, while the potential loss of buying

[2]A put option gives the owner of the option the right, but not the obligation, to sell an asset (the underlying) at a specific price (the strike), by a pre-determined date (the expiration or maturity date) to the seller (or "writer") of the option. A call option gives the buyer of the option the right, but not the obligation, to buy an agreed quantity of the underlying from the seller of the option before the expiration date at a given strike price.

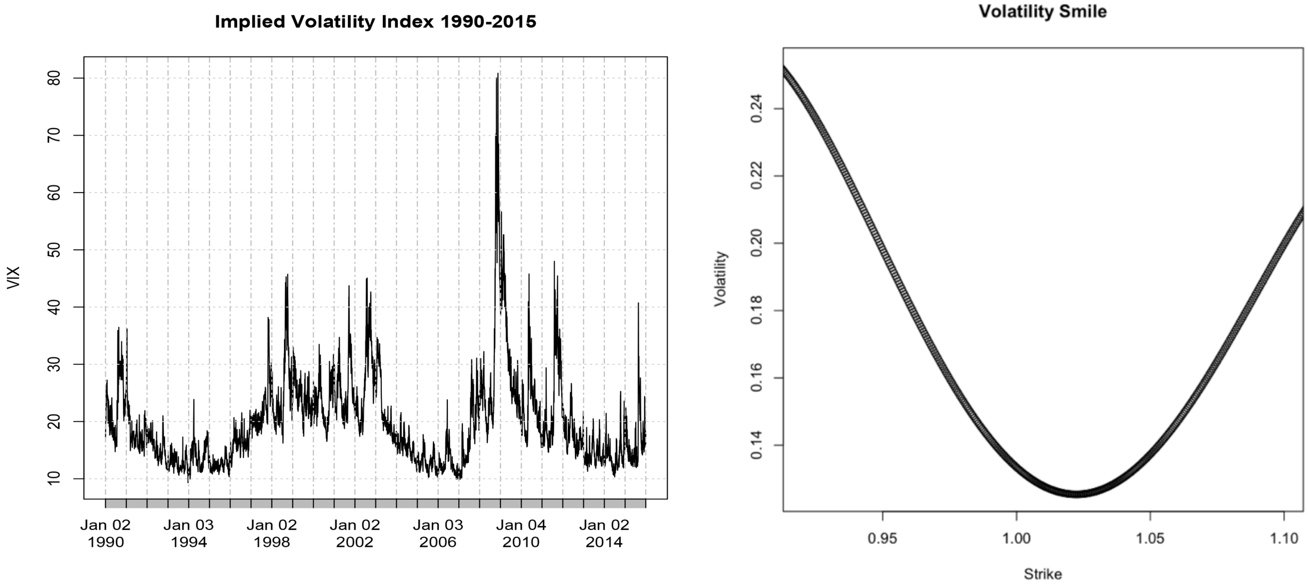


Fig. 4. Implied Volatility Index and the Implied Volatility Smile

options is limited by the premium paid, shorting options could be very risky because the loss is only limited by the market actions. For example, shorting a call option while the market continues moving up could result in a severe loss.

3.2. Options Pricing Model

To derive the fair option price, a common practice is to assume that the underlying asset follows a certain dynamic model such as geometric Brownian motion (GBM) with constant drift and volatility, described by the following stochastic differential equation:

$$dS = \mu S\,dt + \sigma S\,dW \tag{7}$$

where S is the asset price, μ is the drift parameter, σ is the volatility, and W is a Wiener process or Brownian motion. With the assumed model, a closed-form options pricing model has been developed as a function of the current asset price, the option strike price, the time to maturity, and the asset volatility [12, 25]. This popular Back-Scholes-Merton (BSM) option pricing model developed in 1973 had revolutionized the derivative industry for the last several decades.

As mentioned above, an important assumption behind the derivation of the BSM pricing model is that the price of the underlying asset follows a GBM model with constant drift and volatility. However, since the stock market crash of October 1987, the volatility of stock index options implied by the market prices has been observed to be "skewed" in the sense that the volatility became a function of strike and expiration instead of remaining a constant. This phenomenon, referred to as the "volatility smile," has since spread to other markets [26]. Because the original BSM model can no longer account for the smile, investors have to use more complex models to value and hedge their options. In this paper, for the purpose of evaluating the trading performance, we will emulate the option prices subject to the smile phenomenon by utilizing the historical implied volatility index (VIX) data and approximate the volatility smile with a quadratic function of moneyness[3] [25]. Figure 4 shows the historical VIX data from 1990 to 2015 and the corresponding volatility smile based on normalized strike price used in the simulated trading.

3.3. Trading Process

We employ a simple trading strategy, called "strangle," by simultaneously writing both out-of-money[4] (OTM) weekly put and call options. With an expiration cycle of 4 weeks, the options are written repeatedly on a weekly basis. The trading equity is allocated over the 4 weeks period where at the end of each week, a portion of the options expires and a new set of options is initiated/written. We use historical end-of-the-day S&P settlement prices and the options pricing model (Section 3.2) to emulate the filled-prices of the transactions. We assume no transaction cost and no slippage. Note that since the strategy does not produce substantial amount of trading as will be clearly described in the next Section, this assumption does not affect the validity of results.

4. TEST AND SIMULATION

In order to estimate the optimal Kelly's fraction, we first apply the "strangle" strategy with various parameters such as different out of the money (OTM) strikes and maturity dates to test the performance. Specifically, for each set of parameters, the winning probability, the average gain of winning trades, and the average loss of

[3]Moneyness is the relative position of the current price of an underlying asset with respect to the strike price of derivative.

[4]The strike of a call option is above the market price or the strike of a put option is below the market price of the underlying asset.

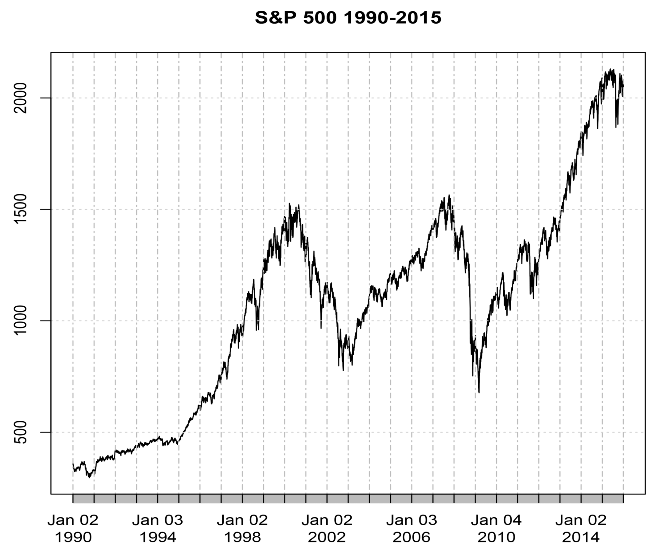


Fig. 5. S&P Index (1990–2015)

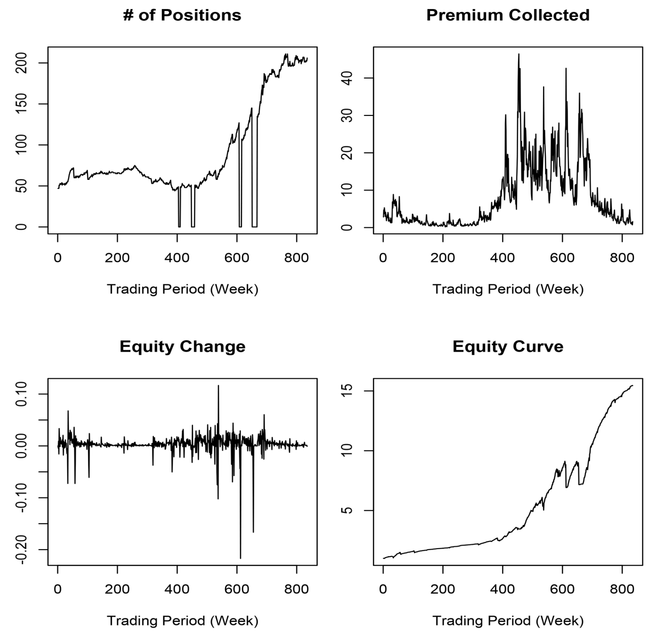


Fig. 6. Options Writing with 6% OTM Strangle (1990–2005)

the losing trades are computed. The results are then used to obtain the optimal Kelly's fraction based on Eqn. (5). The resulting fraction is then applied to allocate equity in the simulated trading process.

4.1. Options Writing Strategy

With the strangle strategy, we simultaneously short the OTM S&P put and call options regardless of the market conditions. We will keep the option positions open until expiration before repeating the same process again in the next trading cycle. To compare the performance, we vary the strike prices of the options from at-the-money (ATM) to 6% OTM with a 1% increment.

Note that the options could expire OTM, and therefore become worthless. In that case, the premium collected by the seller becomes the profit and the positions will be closed automatically by the exchange. On the other hand, if the options expire in the money (ITM), the options will have to be settled in cash in the sense that the sellers have to pay the market price at the expiration time to "buy" back the options they sold. In that case, if the market price deviates more than the premium collected, the seller will incur a loss.

4.2. Simulated Trading

Figure 5 shows the historical S&P data from 1990 to 2015. Since there are only limited historical options prices with specific strikes and expirations available in the public domain, we simulate the options filled-prices based on the model described earlier. Specifically, options prices are obtained by utilizing the BSM model given the S&P price, risk-free interest rate, volatility, and an expiration time of 4 trading weeks after writing the options. The S&P prices are based on historical data and served as the ATM strike prices. Risk-free interest is based on historical 3-month LIBOR data [27] and volatility is based on the historical implied volatility index (VIX).

However, as mentioned earlier, it is well known that true volatility is not a constant but a function of strike and expiration (volatility smile and surface, see Figure 4) [28]. To obtain a more realistic options price, we develop a smile model and adjust the option price accordingly as described in Section 3.2. The results have been validated against the available market data and proved to be reasonably accurate.

4.3. Performance Results

To estimate the Kelly parameters, we use 16 years (1990–2005) of historical data to test the weekly strangle options writing strategy. To be conservative, a 30% margin[5] is assumed to be required for each option contract. In addition, a VIX threshold of 35 is set to avoid a potential catastrophic loss.[6] In other words, all position will be closed when the VIX goes beyond the threshold and new positions will not be written until VIX moves below the threshold. To test the performance, the strike prices are varied from 0% ATM to 6% OTM. For example, Figure 6 shows that with 6% strangle writing, the option strategy produces fairly smooth equity curve with some minor drawdowns. Note that the top left panel of Figure 6 shows that the number of positions drop to zero in several occasions [20]. This is due to the VIX based closing criterion mentioned above.

The detailed resulting performances are summarized in Figure 7. For example, with a 6% OTM strike, the rate of winning is around 80% and the draw down is about 24%. The average amount of winning[7] is 2.2

[5]Margin is the amount of capital needed to initiate and maintain an option position. Typically, for S&P options, the margin requirement could be as low as 10% of the underlying asset value.

[6]The historical average of VIX is below 20.

[7]Note that for e-mini S&P futures market, each point corresponds to $50.

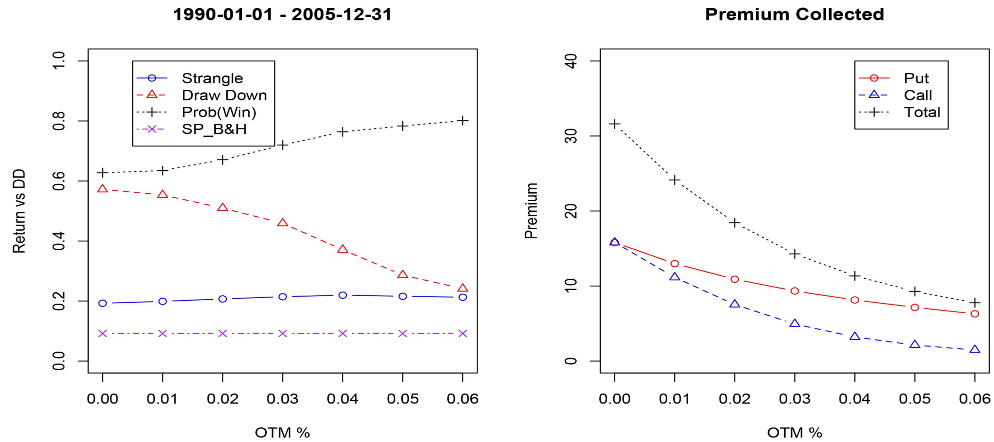


Fig. 7. Performance Summary—Strangle Writing (1990–2005)

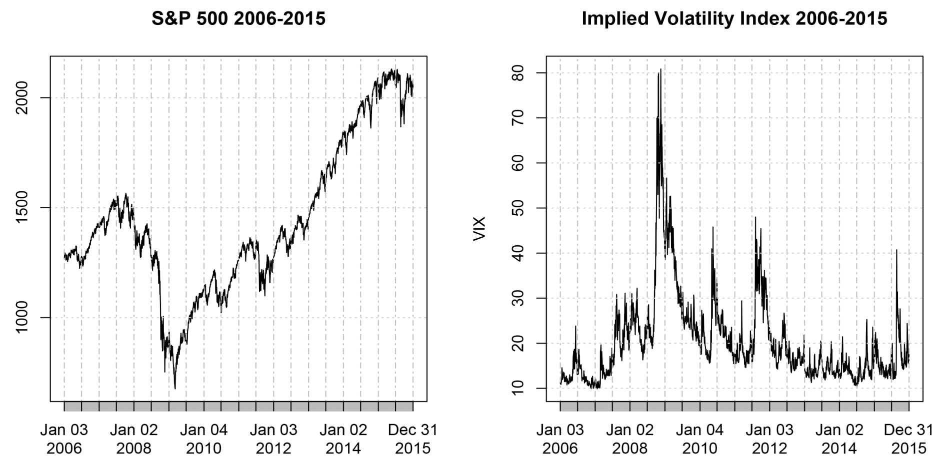


Fig. 8. S&P and VIX—2006–2015

per winning trade and −4.2 for a losing trade. This corresponds to a Kelly's fraction of

$$f^* = \frac{bp - (1-p)}{b} = \frac{\frac{2.2}{4.2}(0.8) - 0.2}{2.2/4.2} \approx 0.418. \quad (8)$$

The results indicate that, based on the historical performance, the Kelly's formula recommends an allocation of about 42% of the equity for each trade in order to achieve the highest possible long-term gain. Note that Figure 7 also shows that the average S&P annual return (Buy and hold) during the same period was less than 9%.

We apply the Kelly criterion to test the S&P data from 2006 to 2015. Figure 8 shows the corresponding S&P and VIX data over the 10-year period. Note that a 42% Kelly also implies that at most 42% of the equity can be lost in a single trade. In order to ensure that, a stop loss needs to be in place to determine the total number of option positions that could be written. For example, with an initial capital of $1M, a 42% Kelly and a $4k stop loss per contract, the maximum number of positions is $1M*0.42/$4k = 105.

Figure 9 shows the trading performance with a 6% strangle and the optimal Kelly's fraction. The option strategy produces an average of 40+% annual return and a draw down of around 40%. The detailed performances are shown in Figures 10–11 and also summarized in Table 1. In the table, two sets of performance results, one with the optimal Kelly and the other with a $\frac{1}{2}$ Kelly, are given for comparison. It can be seen from the table that, with the optimal Kelly, the strangle strategy generally produces much higher rate of return than the naïve buy-and-hold (B&H) policy, at the expense of a higher risk (DD) [20].

For example, Table 1 shows that over the 10-year period, the B&H strategy produces an average annual return of 5.73% with a DD of 56.24%, while a 2% OTM strangle produces a 26.39% annual return with a DD of 74.56%. On the other hand, with a 5% OTM strangle, the annual return reach the highest value of 44.82% with a draw down of 50.02%. This relatively "conservative" strategy[8] produces much better return than the naïve B&H strategy while with a smaller drawdown.

[8]A strangle selling strategy with higher OTM strikes is more conservative than the one with lower OTM strikes in the sense that it is expected to achieve a lower rate of return with a smaller DD.

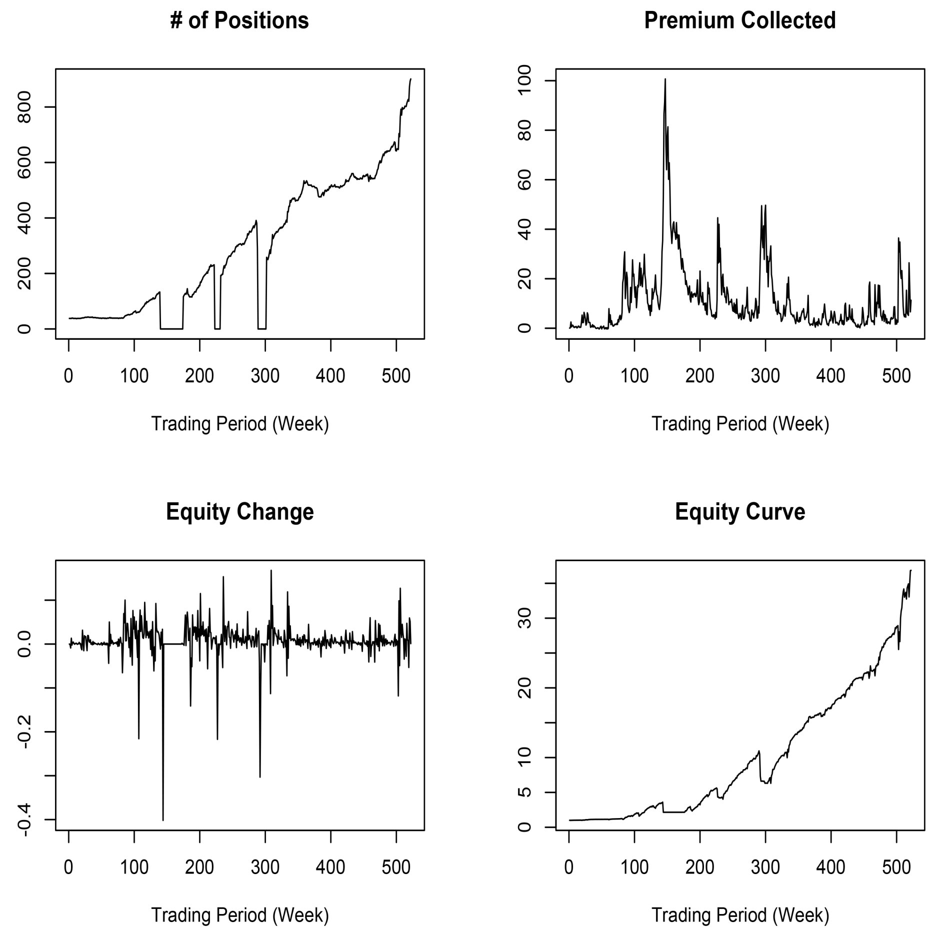


Fig. 9. Options Writing Performance with 6% OTM Strangle and Optimal Kelly Fraction (2006–2015)

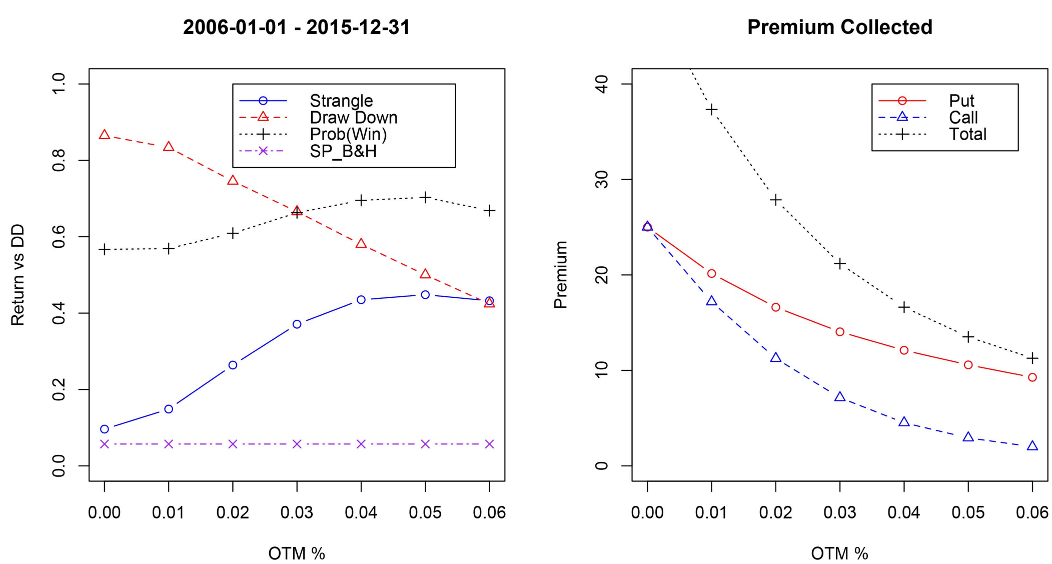


Fig. 10. Performance Summary—Strangle Writing with Optimal Kelly Fraction (2006–2015)

With $\frac{1}{2}$ Kelly, both annual rate of return and maximum drawdown are much lower due to lower leverage as shown in Figure 11 and Table 1. For example, with a 4% strangle, the annual rate of return reduces to 23.28% while the draw down also drops to 30.75%. A 6% strangle reduces the DD to 22.48% and an average annual return of 21.22%. It is clear from the table that by selecting different leveraged options writing policies, the performance can be adapted to fit individual investor's risk aptitude.

Figure 12 shows the trade-off between risk and return with different OTM strangles and different invest-

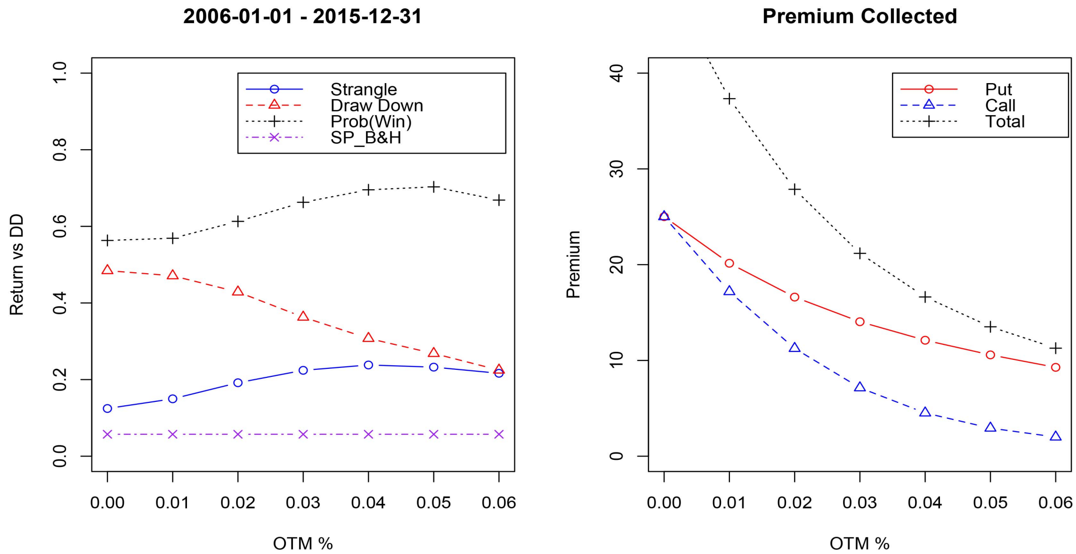


Fig. 11. Performance Summary—Strangle Writing with $\frac{1}{2}$ Kelly fraction (2006–2015)

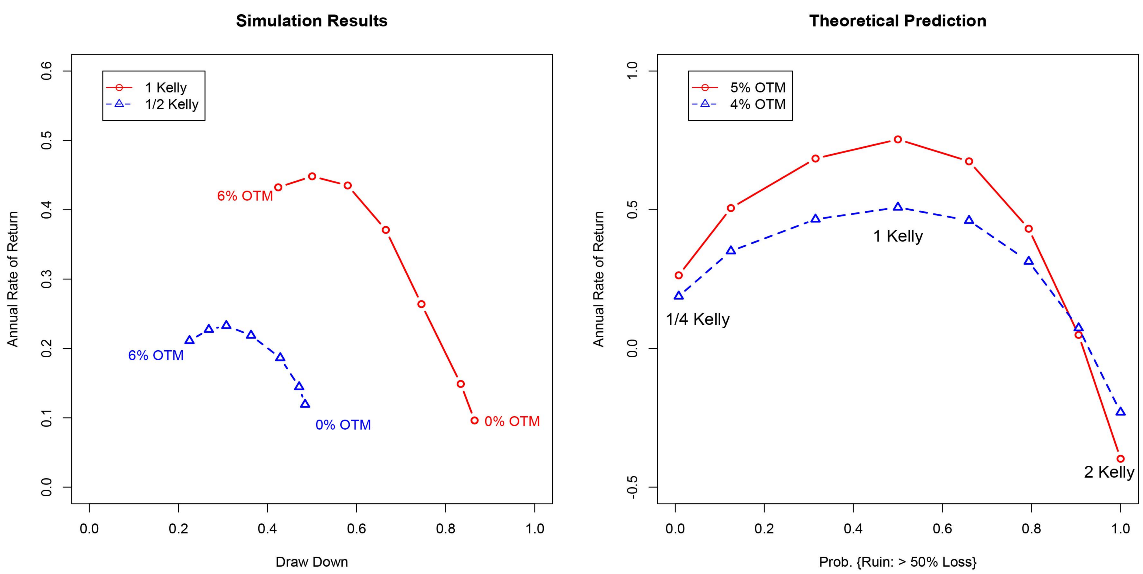


Fig. 12. Performance Summary vs. Theoretical Predictions

TABLE 1.
Performance Comparison with optimal Kelly fraction

2006–2015	Annual RR 1 Kelly	Max DD 1 Kelly	Annual RR $\frac{1}{2}$ Kelly	Max DD $\frac{1}{2}$ Kelly
Buy–Hold	5.73%	56.24%	5.73%	56.24%
Strangle 0%	9.62%	86.49%	11.92%	48.44%
Strangle 1%	14.88%	83.38%	14.45%	47.10%
Strangle 2%	26.39%	74.56%	18.64%	42.86%
Strangle 3%	37.09%	66.56%	21.88%	36.30%
Strangle 4%	43.50%	58.00%	23.28%	30.75%
Strangle 5%	44.82%	50.02%	22.72%	26.82%
Strangle 6%	43.23%	42.41%	21.11%	22.48%

ment fractions [20]. The left panel of Figure 12 shows that, for the simulated trading during the 10-year period, the highest rate of return could be achieved with around 4–5% OTM strangles regardless of the choice of Kelly's fractions. The right panel shows the theoretical prediction using Equations (4) and (6) based on the probability of winning and average gain and loss per trade obtained from the simulation results for 4% and 5% OTM strangle options writing. As expected, the "efficient frontier" peak at the optimal Kelly's fraction. With the SOC curves given in Figure 12, an operating point can be chosen to satisfy almost any desired risk aptitude, should that be defined as drawdown or probability of ruin. For example, an aggressive investor might decide to employ a higher Kelly leverage ratio and a higher OTM strike price with an expectation of better return and an understanding of the accompanying higher risk as indicated by the predictions.

5. MULTI-ASSET ALLOCATION

With the "mean-variance" (MV) approach described in Section 2.1, an efficient frontier can be constructed for multi-asset allocation where for each return level a

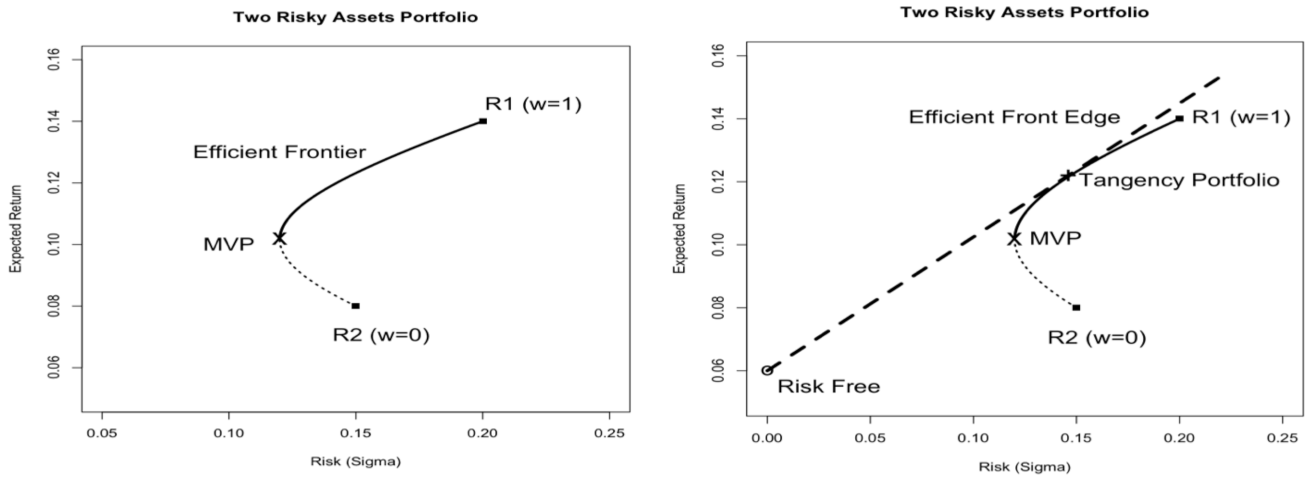


Fig. 13. (a) Efficient Frontier for a Two Risky Asset Portfolio; (b) Efficient Front Edge and Tangency Portfolio

portfolio can be derived with minimum risk (variance). For example, with two risky assets, $\boldsymbol{\mu} = [0.14, 0.08]^T$ and $\Sigma = \mathrm{diag}[0.2^2, 0.15^2]$, Figure 13 (a) shows the "efficient frontier" portfolios on a mean-STD (standard deviation) chart. Each optimal portfolio consists of a combination of the two assets where at the top right end of the frontier, the entire 100% of the capital is allocated to asset 1 while at the bottom, all capital is allocated to asset 2. Any portfolio below the return level of MVP (minimum variance portfolio) is not considered efficient due to its lower expected return. Therefore, an efficient frontier is constructed from MVP to the top right of the curve.

With the efficient frontier, an investor could choose a portfolio on the curve depending on his/her own risk aptitude. For example, a conservative investor may choose a portfolio close to MVP while an aggressive investor may choose a portfolio close to R_1.[9] Note that when risk-free asset[10] is available, an "efficient front edge" can be constructed by connecting the risk-free asset and the tangency portfolio on the mean-variance chart[11] (see Figure 13(b) with risk-free rate $r_f = 0.06$). The MV tangency portfolio (MVTP) can be shown to maximize the risk-adjusted return (Sharpe ratio[12]) and is a desirable choice of optimal portfolio on the frontier[13] [9].

5.1. Kelly's Approach

As in (2), we consider a portfolio consisting of a set of correlated assets with weights $\boldsymbol{\omega} = [\omega_1 \cdots \omega_N]^T$ such

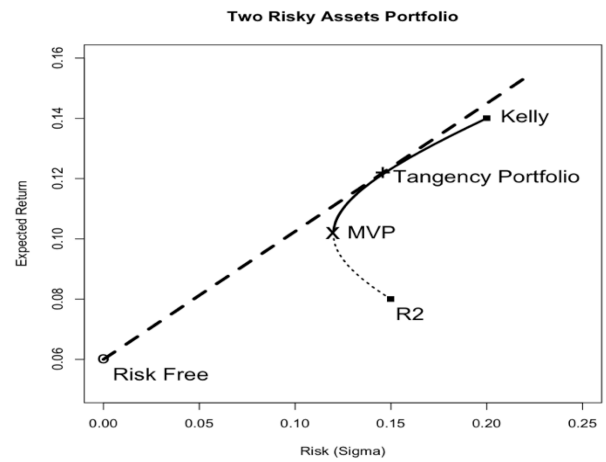


Fig. 14. Tangency Portfolio and the Kelly's Portfolio

that $\sum_{i=1}^{N} \omega_i = \mathbf{1}^T\boldsymbol{\omega} - 1$. Then the portfolio value for the following investment period becomes,

$$P(k) = P(k-1)\left(1 + \sum_{i=1}^{N} \omega_i R_i(k)\right) = P(k-1)(1 + R_P(k)) \quad (9)$$

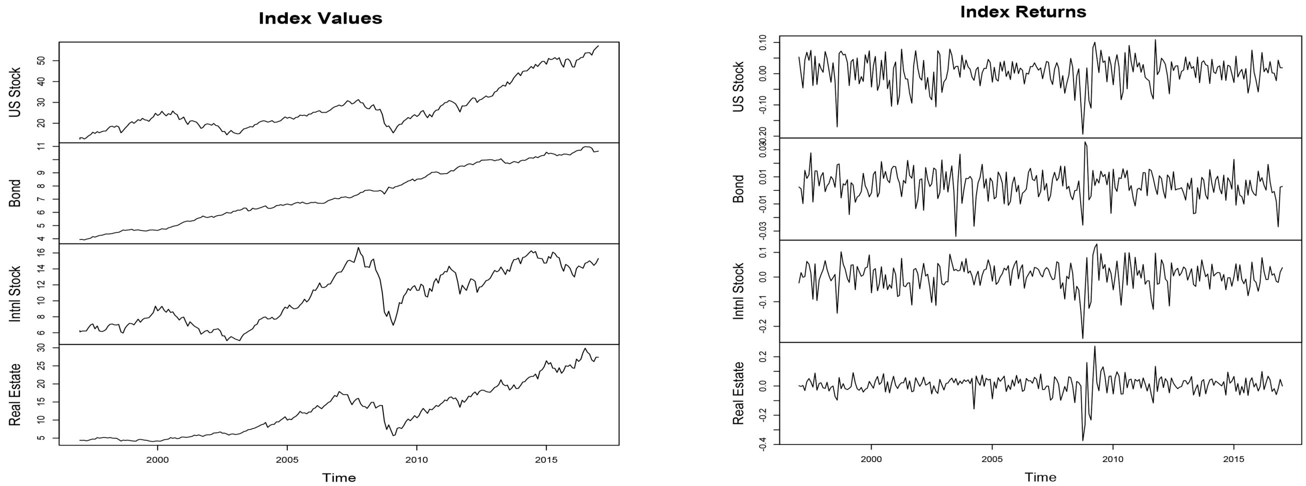
where $P(k)$ is the portfolio value at time k and $R_P(k)$ is the portfolio return.

According to Kelly, in order to maximize the investment growth rate in the long run, it is equivalent to maximize the logarithm of the equity after each time step [13, 29–30]. Therefore, to construct Kelly's portfolio, with no short selling and no leverage, it is necessary to solve the following optimization problem,

$$\mathrm{Max}_{\omega}\, E\left[\ln\left(1 + \sum_{i=1}^{N} \omega_i R_i(k)\right)\right] \text{ subject to } \boldsymbol{\omega}^T\mathbf{1} = 1; \omega_i \geq 0 \quad (10)$$

For example, the Kelly's portfolio (KP) for the two risky assets example given in Figure 13 turns out to be the one with 100% allocation on asset #1 as shown in Figure 14.

[9] If short-selling or borrowing/leverage are allowed, an investor could choose a portfolio beyond R_1 where higher expected return together with higher risk can be achieved.
[10] Cash or Treasuries with interest rate r_f and with little or no risk.
[11] It's also called the capital market line (CML).
[12] Sharpe ratio is the risk-adjusted return defined as $(\mu_P - r_f)/\sigma_P$
[13] When the portfolio includes all assets in the market, the tangency portfolio converges to the market portfolio by the equilibrium argument [11].



(a) (b)

Fig. 15. (a) The Four Core Asset Index Funds: 1996–2016; (b) Core Asset Returns: 1996–2016

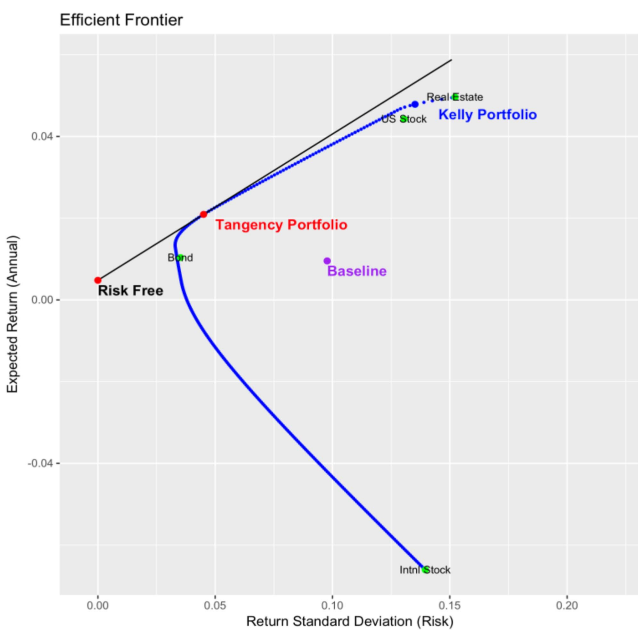


Fig. 16. Efficient Frontier and the Kelly's Portfolio

5.2. Test and Simulation

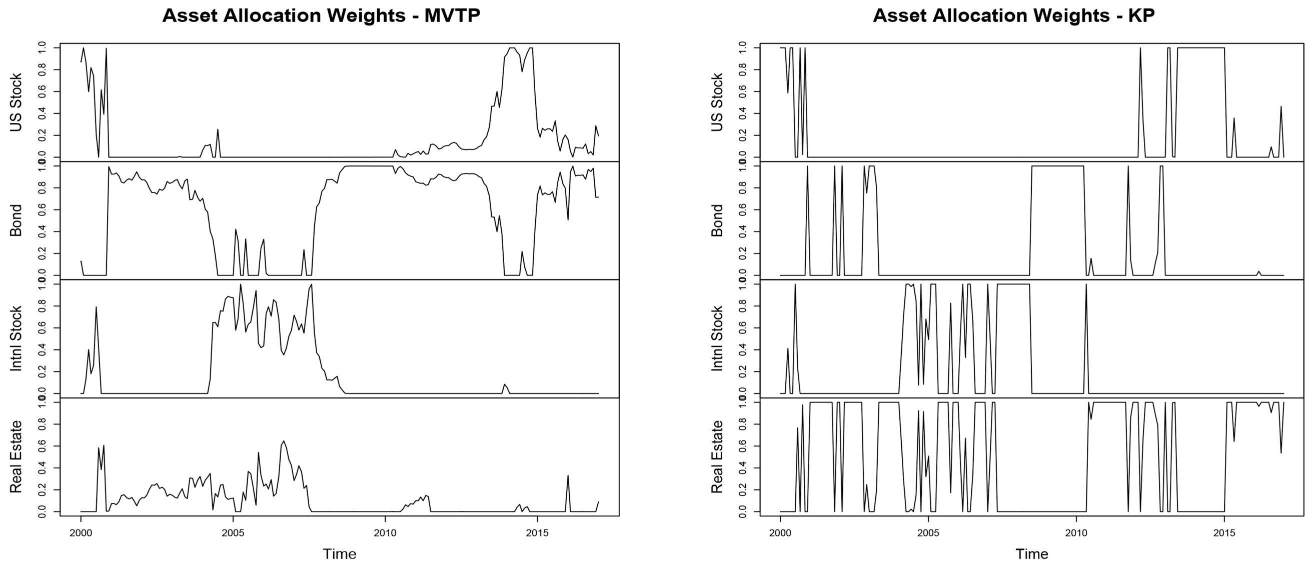
For evaluation purpose, we compose an artificial portfolio with a few "core assets" selected from the Vanguard index-based mutual fund family. Specifically, four core asset classes, including the US total stock market index (VTSMX), total bond index (VBMFX), international stock index (VGTSX), and the real estate index (VGSIX), are selected to construct the portfolio. The 20-year historical prices obtained from finance.yahoo.com and their log returns of the four assets from 1996 to 2016 are shown in Figure 15.

As mentioned earlier, the goal of constructing an optimal portfolio is to maximize the investor's return or minimizing the risk. Under a given capital constraint, portfolios are constructed and dynamically rebalanced by allocating the capital over the four core assets using different strategies. With no shorting and no leverage assumptions, the allocation fraction of each asset is subject to $\boldsymbol{\omega}^T\mathbf{1} = 1$ and $\omega_i \geq 0$.

The three strategies to be compared include MV tangency portfolio (MVTP), Kelly's portfolio (KP), and the portfolio based on a Naïve strategy. The Naïve strategy is a simple allocation scheme served as the baseline for comparison, in which the portfolio is simply rebalanced uniformly among all the core assets at the beginning of each investment period. The historical data of the four core assets are used to train the model. Specifically, a sliding window of 18 months of data is used to estimate the asset returns, volatilities, and the correlations between the assets. Based on the results, optimal portfolios under each strategy will be formed and rebalanced accordingly on a monthly basis from 2000 to 2016. During the test period (2000–2016), a total of 204 months is available for portfolio rebalancing and performance evaluation. For example, Figure 16 shows a snap shot of the efficiency frontier and the corresponding locations of the three strategies for Nov. 2016. The history of the dynamic allocation fractions of the four core assets in the portfolio based on MVTP and KP are shown in Figure 17. As can be seen, KP tends to take a more extreme allocation than that of the MVTP.

5.3. Performance Results

With the three strategies, the portfolio is dynamically rebalanced monthly during the investment period from 2000 to 2016. In the process, we assume no transaction cost and no slippage for the rebalancing trades. The resulting portfolio equity curves for the three strategies are shown in Figure 18 and the overall performances are summarized in Table 2. The Naïve portfolio has the lowest annualized return (5.33%) with highest volatility (12.80%) and worst drawdown (49.94%) due to its simplicity and the inability to deal with the 2008



(a) (b)

Fig. 17. (a) Asset Allocation Fractions—MVTP; (b) Asset Allocation Fractions—KP

TABLE 2.
Performance Summary: 2000–2016

2000–2016	MVTP	Baseline	KP
Annual Return	7.44%	5.33%	**9.90%**
Annual Risk	**6.65%**	12.80%	12.53%
Sharpe Ratio	**1.044**	0.377	0.750
Drawdown	**14.78%**	49.94%	23.27%

credit crisis. As expected, Kelly's portfolio (KP) is an aggressive strategy and it produces the highest annualized return (9.90%) while suffers a noticeable drawdown (23.27%). Not surprisingly, the MVTP strategy produces the highest Sharpe ratio of 1.044. This is expected by the nature of the tangency portfolio. In addition, it's necessary to point out that the MVTP is able to weather the 2008 credit storm with a fairly small portfolio volatility (6.65%) and a manageable drawdown (14.78%). Note that MVTP and KP represent two complementary strategies that allow an investor to make a tradeoff between risk and return according to his/her own preferences.

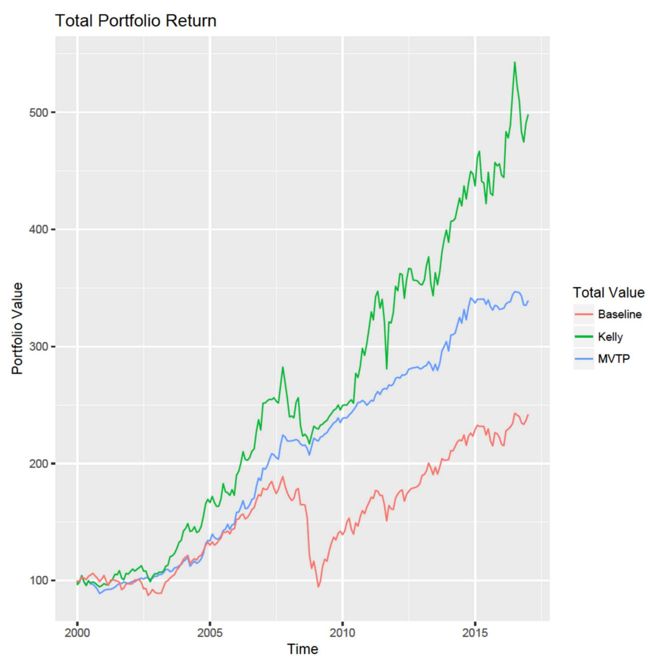


Fig. 18. Equity Curves Comparison

6. SUMMARY AND CONCLUSION

An optimal asset allocation strategy to support investment and trading decisions is developed. First, a simple yet practical trading scenario where the entire equity is allocated between a risk-free asset and a risky asset is considered. The goal is to identify the optimal fraction of the equity to be allocated for trading in order to achieve the highest long-term return with a limited risk. The allocation is based on Kelly criterion derived from the concept of mutual information in binary symmetric communication channels. The resulting model is applied to dynamically allocate equity for writing S&P futures options. Several trading strategies are implemented based on the decision makers' risk aptitude.

Similar to the classical portfolio theory, a system operating curve is developed for each trading strategy where each operating point on the curve representing an expected trade-off between risk of ruin and return. An investor can choose any operating point to satisfy a desired risk and return aptitude. An extended simulation was conducted for performance prediction and evaluation by incorporating historical market data of S&P index futures and options. The results of the simulated trading using these strategies over a 10-year period significantly outperform the buy-and-hold strategy. They are also consistent with the analytical performance predictions.

The single asset allocation strategy is then extended to portfolio rebalancing with multiple correlated assets. As in the single asset case, the goal is to maximize the long-run portfolio growth rate over many investment cycles. We compare three strategies including MV tangency portfolio (MVTP), Kelly's portfolio (KP), and the Naïve strategy. Through extensive out-of-sample simulations, we show that the resulting KP strategy outperforms the traditional MVTP approach in annual return but with higher volatility. As expected, KP can be considered as an alternative approach for investor to trade off between risk and reward.

While the preliminary results shown in this paper are promising, one of the critical future step is to develop and integrate a dynamic model [31–32] into the allocation strategies so that we would be able to apply the expected future returns of the chosen assets into the optimization process. Another potential future research direction is to integrate the quantitative data with the qualitative information by utilizing the data fusion and machine learning technologies.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Jiayang Yu for his input and programming assistance on multi-asset allocation.

REFERENCES

[1] S. Martin
Macroeconomic Variables and Stock Market: US Review
MPRA Paper No. 39094, May 2012. http://mpra.ub.uni-muenchen.de/39094/.

[2] X. P. S. Zhang and F. Wang
Signal Processing for Finance, Economics, and Marketing: Concepts, framework, and big data applications
IEEE Signal Processing Magazine, vol. 34, no. 3, pp. 14–35, May 2017.

[3] Y. Zuo and E. Kita
Stock Price Forecast using Bayesian Network
Expert Systems with Applications, Elsvier, Vol. 39, pp. 6729–6737, 2012.

[4] O. Jangmin, J. W. Lee, S. B. Lark, and B. T. Zhang
Stock trading by Modeling Price Trend with Dynamic Bayesian Networks
IDEAL, pp. 794–799, 2004.

[5] J. Geweke and G. Amisano
Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns
European Central Bank Eurosystem, Working paper No. 969, 2008.

[6] N. G. Polson and B. V. Tew
Bayesian Portfolio Selection: An Empirical Analysis of the S&P 500 Index 1970–1996
ASA Journal of Business and Economic Statistics, Vol. 18, No. 2, 2000.

[7] R. Demirer, R. R. Mau, and C. Shenoy
Bayesian Networks: A Decision tool to Improve Portfolio Risk Analysis
Journal of Applied Finance (Winter 2006), 106–119.

[8] C. Shenoy and P. P. Shenoy
Bayesian Network Models of Portfolio Risk and Return
Computational Finance 1999, pp. 87–106, The MIT Press, Cambridge, MA.

[9] J. Hull
Options, Futures, and other Derivatives, 9th edition
Pearson, 2015.

[10] I. Karatzas and S. Shreve
Methods of Mathematical Finance
Springer-Verlag New York, Incorporated, 1998.

[11] C. Cao and J. Hiang
Determinants of S&P Index Option Returns
Review Derivative Research, 10:1–38, 2007.

[12] F. Black and M. Scholes
The pricing of options and corporate liabilities
Journal of Political Economy 81, 637–653, 1973.

[13] J. Kelly
A New Interpretation of Information Rate
Bell System Technical Journal 35: 917–926, 1956.

[14] E. Thorp
Fortune's Formula: The Game of Blackjack
American Mathematical Society, Jan, 1961.

[15] E. Thorp
The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market
http://www.eecs.harvard.edu/cs286r/courses/fall12/papers/Thorpe_KellyCriterion2007.pdf.

[16] E. Thorp
A Man for All Markets: From Las Vegas to Wall Street, How I beat the Dealer and the Market
Random House, 2017.

[17] E. Platen and. Heath
A Benchmark Approach to Quantitative Finance
Springer, 2006.

[18] T. M. Cover and J. A. Thomas
Elements of Information Theory, 2nd edition
Wiley, 2006.

[19] Y. Feng
Dynamic Asset Allocation; A Bayesian Approach
Ph.D. Dissertation, City University of New York, 2014.

[20] K. C. Chang and Z. Tian
Optimal Asset Allocation with Mutual Information
Proc. 19th International Conference on Information Fusion, July 5–8, 2016, Heidelberg, Germany.

[21] H. Markowitz
Portfolio Selection
The Journal of Finance, Vol. 7, p. 77–91, 1952.

[22] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann
Modern Portfolio Theory and investment Analysis, 9th edition
Wiley, 2014.

[23] CME
http://www.cmegroup.com/trading/equity-index/us-index/sandp-500_contract_specifications.html.

[24] Globex
http://www.cmegroup.com/trading/equity-index/us-index/e-mini-sandp500.html.

[25] R. C. Merton
Theory of Rational Option Pricing
Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1): 141–183, 1973.

[26] D. Backus, S. Foresi, and L. Wu
Accounting for Bias in Black-Scholes
http://faculty.baruch.cuny.edu/lwu/papers/bias.pdf.

[27] Federal Reserve Economic Data, St. Louis FRED
http://research.stlouisfed.org/fred2/.

[28] J. D. Coval and T. Shumway
Expected Option Returns
The Journal of Finance, Vol. LVI, No. 3, 2001.

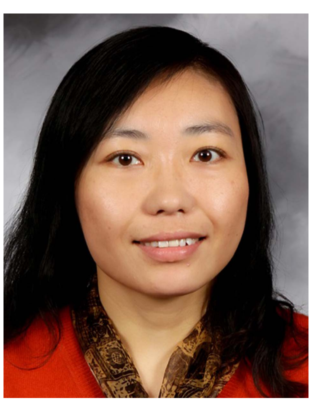
[29] P. Laureti, M. Medo, and Y. Zhang
Analysis of Kelly-optimal Portfolios
Quantitative Finance, Vol. 10, No. 7, p. 689–697, Aug., 2010.

[30] V. Nekrasov
Kelly criterion for multivariate portfolio: a model-free approach
Elsevier, SSRN, 2014. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2259133.

[31] K. Triantafyllopoulos
Time-varying vector autoregressive models with stochastic volatility
Journal of Applied Statistics, Vol. 38, No. 2, p, 369–382, Feb., 2011.

[32] M. West and J. Harrison
Bayesian Forecasting and Dynamic Modes
Springer, 2nd edition, 1997.

Kuo-Chu Chang received the M.S. and Ph.D. degrees in Electrical Engineering from the University of Connecticut in 1983 and 1986 respectively. From 1983 to 1992, he was a senior research scientist in Advanced Decision Systems (ADS) division, Booz-Allen & Hamilton, Mountain View, California. In 1992, he joined the Systems Engineering and Operations Research department, George Mason University where he is currently a professor. His research interests include estimation theory, multisensor data fusion, Bayesian inference, and financial engineering. He is particularly interested in applying unconventional techniques in the conventional decision and control systems. He has more than 30 years of industrial and academic experience and published more than two hundred papers in the areas of multitarget tracking, distributed sensor fusion, Bayesian Networks technologies, and financial engineering. He was an associate editor on Tracking/Navigation Systems from 1993 to 1996 and on Large Scale Systems from 1996 to 2006 for *IEEE Transactions on Aerospace and Electronic Systems*. He was also an associate editor of *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, from 2002 to 2007. He was the Technical Program Co-chair of the 2009 International Conference on Information Fusion, Seattle, USA.



Zhi Tian is a Professor in the Electrical and Computer Engineering Department of George Mason University, since 2015. Prior to that, she was on the faculty of Michigan Technological University from 2000 to 2014. Her research interests lie in estimation and detection theory, with applications in statistical signal processing, wireless communications, and target tracking and inference. She is an IEEE Fellow. She serves as an IEEE Distinguished Lecturer for both the IEEE Communications Society and the IEEE Vehicular Technology Society. She served as Associate Editor for *IEEE Transactions on Wireless Communications* and *IEEE Transactions on Signal Processing*.