UNIVERSIDAD DEL CEMA
Buenos Aires
Argentina

Serie
DOCUMENTOS DE TRABAJO

Área: Finanzas

MEASURING AND TRADING VOLATILITY ON THE US STOCK MARKET: A REGIME SWITCHING APPROACH

José P. Dapena, Juan A. Serur y Julián R. Siri

Septiembre 2018
Nro. 659

www.cema.edu.ar/publicaciones/doc_trabajo.html
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding jae@cema.edu.ar

Measuring and trading volatility on the US stock market: A regime switching approach

July, 2018

José P. Dapena[1], Juan A. Serur[2] and Julián R. Siri[3]

Universidad del CEMA
Av. Córdoba 374, C1054AAP, Ciudad de Buenos Aires, Argentina

Abstract

The volatility premium is a well-documented phenomenon, which can be approximated by the difference between the previous month level of the VIX Index and the rolling 30-day close-to-close volatility. Along with the literature, we show evidence that VIX is generally above the 30-day rolling volatility giving rise to the volatility premium, so selling volatility can become a profitable trading strategy as long as proper risk management is under place. As a contribution, we introduced the implementation of a Hidden Markov Model (HMM), identifying two states of the nature and showing that the volatility premium undergoes temporal breaks in its behavior. Based on this, we formulate a trading strategy by selling volatility and switching to medium-term U.S. Treasury Bills when appropriated. We test the performance of the strategy using the conventional Carhart four-factor model showing a positive and statistically significant alpha.

JEL: C1, C3, N2, G11.
Key words: Realized volatility, expected volatility, volatility premium, regime switching, excess returns, hidden markov model, VIX.

[1] José P. Dapena, Ph.D. Economics, is the Director of the Department of finance at University of CEMA. Email: jd@ucema.edu.ar

[2] Juan A. Serur, M. Finance, is an Assistant Professor at University of CEMA. Email: juanandresserur@gmail.com

[3] Julián R. Siri, MSc. Financial Engineering, is a Professor at University of CEMA. Email: julian.siri@gmail.com

We sincerely appreciate the comments and suggestions made at the IV Jornadas Argentinas de Econometría. The author's views are on their own and do not necessarily represent those of the Universidad del CEMA.

Introduction

Selling insurance can become a profitable strategy as long as the appropriate risk management methodology is under place. Finance is about risk, and risk is about uncertainty. In this sense, models in finance have been historically aimed at capturing and appropriately measuring risk and evaluating the appropriate expected returns that compensates for it. Investors can act as insurance providers by selling risk when they consider it is expensive, i.e., expected returns do not compensate for it, and buy it when it is supposedly cheap, i.e., expected returns overcompensate for risk. Hence a model that helps identifying when risk is expensive is of much value. In most of the cases, the risk is approximated by the volatility of returns, under the assumption that investors hold certain utility functions, or that returns follow a normal distribution.[4]

The financial contracts that capture risks very well are given by options. In fact, the basic option pricing model aims at eliminating the drift from the calculation to put a price on volatility. Considering the importance of the volatility in finance, in 1993 the Chicago Board Exchange (CBOE) introduced the Volatility Index (VIX Index), which became one of the main USA capital market indicators. The investor community labeled this index as the "uncertainty index" or "fear gauge index", intended to show market sentiment about uncertainty[5]. In 2004 and 2006, futures and options on the VIX Index were respectively introduced, and from this point in time, uncountable VIX-related products were put in place. The VIX Index captures the implied 30-day volatility of the S&P 500 Index through option prices, becoming an index which reflects market expectation about future uncertainty. Given that VIX is an index about expected risk or volatility, its value can be then compared with realized volatility to gauge how well the market predictions work. Studies show that the implied-realized spread for the S&P 500's volatility can be to some extent approximated by the spread between the rolling 30-day close-to-close realized volatility and the previous level of VIX, showing that VIX is generally above the 30-day rolling volatility. This results in what is known as "volatility premium"[6].

It is the purpose of the present work to formally test the volatility premium through a quantitative model with the intention to exploit the fact that sometimes volatility is expensive (e.g., implied volatility is greater than realized volatility) and hence it becomes profitable to sell it. We show how to exploit this phenomenon by using different volatility vehicles that aim at gaining exposure to the underlying in

[4] However, more advanced models of finance aim to capture superior moments of the distribution (see, [Kraus and Litzenberger, 1976]), and incorporate skewness and coskewness (see, e.g., [Harvey and Siddique, 2000] for an extension of the CAPM model to three moments).

[5] It is important to mention that by uncertainty we refer to the risk measured by VIX Index and not to the Knightian uncertainty. Both terms imply something different. For details on the index construction visit, https://www.cboe.com/micro/vix/vixwhite.pdf.

[6] For more details on this phenomenon and pertinent information, see, e.g., [Sinclair, 2013].

different ways by means of a Hidden Markov Model (HMM) where two states of the nature (expensive volatility or high spread on the one hand; and cheap volatility or low spread on the other hand) are identified giving rise to a trading strategy.

As for the paper, the structure proceeds in the following way: in Section 1 we introduce the volatility strategies and review the main contributions of the literature to the topic under analysis; in Section 2 we describe thoroughly the implemented methodology; in Section 3 we expose the obtained results; in Section 4.1 we discuss the February sell-off event and its implication to our model. Finally, we discuss the main conclusions.

1 Review and past literature

The price of risk tends to be expensive. In that regards, some authors have shown that there is a profit opportunity from selling protection against the skewness (see, for example, [Ilmanen, 2012], [Ilmanen, 2013]). An investor can sell systematically index options and obtain excess returns (see, for example, [Dapena and Siri, 2015], [Dapena and Siri, 2016]), being consistent with the literature which reports that option buyers tend to earn less returns than predicted by standard risk models.

The difference between expected volatility (expected uncertainty) and realized volatility (what actually happens) has become easier to evaluate since the introduction in 1993 of the VIX Index by the CBOE. The evidence, as mentioned in the previous section, shows that there is a spread between expected volatility and realized volatility, letting think that options of the S&P 500 Index may tend to be overpriced relative to a supposedly "fair price" given by most of the traditional option pricing models, meaning that investor overpays for these financial assets. That could be a reason which explains why the VIX Index, which it is seen by almost all investor as a forward-looking measure of the future realized volatility, tends to consistently overestimate it.

In [Sinclair, 2013], the author mentions that implied (expected)[7] volatility is, in general, an upward biased estimate of realized volatility, where it is common to see that realized volatility is 30% lower than the implied volatility, but practically never see the other way round.[8] He quotes a number of reasons, such us the risk premium associated with selling insurance due to events that never happened but nothing prevents them from happening in the future; that market microstructure encourages implied volatility to be biased high given that market makers are buying insurance, etc. But the pattern in the data does not stop there. VIX futures tend

[7] Expected and implied mean the same for the purpose of the paper.

[8] He also prevents from drawing the incorrect conclusion that there is a sure profit from selling implied volatility.

to overestimate the future value of the spot value of the VIX Index.

This type of events leads to the phenomenon mentioned above, widely known as the “volatility risk premium". Since investors seek for protection and are reluctant to the exposure to downside tail risk, acting as an insurance provider could become profitable because of the willingness to pay. As mentioned, the excess return can be attributed to the insurance premium that investors are willing to pay to obtain protection from events that have never happened.

Among the studies that document the phenomenon, some authors showed that the negative volatility risk premium present in index options have power explanation on the positive difference between implied and realized volatility (see, [Jackwerth and Rubinstein, 1996]). Others showed than on average, Black-Scholes implied volatilities of closest to at-the-money equity options tend to be higher than historical realized volatility (see, [Bakshi and Kapadia, 2003]). Besides, the authors demonstrated that long delta-hedge strategies generate on average statistically significant negative returns, concluding that buyers of volatility leave money on the table. These effects were more pronounced on index options than on single equity options. More evidence presented, showed that the implied volatility is systematically higher than the realized volatility, being on average 19% the first and 16% the second. This difference translates into excess returns for index option sellers (see, for example, [Eraker, 2007])[9]

As mentioned above, many explanations have been proposed about the volatility risk premium. One of the most common conclusions of researchers is that market participants tend to pay more for these derivatives due to the probabilities that they assign to the occurrence of “extreme movements”. On the opposite side, the sellers of volatility require a higher premium for taking the risk of these extreme events. This situation is analogous to the insurance industry on the real economy. One way to exploit all these findings is by mean of volatility-related products like ETNs.

There are main two groups of volatility-related ETNs. On one hand, products with positive exposure to the underlying, and on the other hand products with negative exposure to it. For instance, the VXX ETN provides exposure to the short term VIX futures by daily rebalancing the position between the first (selling) and the second-month (buying) VIX futures. Similarly the VXZ, offers exposure to the medium-term VIX futures.

It is important to mention that the returns of the products with positive exposure to the VIX have been on average negative since their inception. One of the main causes is the shape of the futures curve, which is almost all the time in “contango”,

[9] For more literature and discussion, see, for example, [Andersen, Fusari and Todorov, 2015], [Fleming, 1999], [Kozhan, Neuberger and Schneider, 2013], [Lin and Chen, 2009].

meaning that long-term futures are higher than short-term futures, implying a roll cost due to the daily rebalance, translating to daily losses.

As for products with negative exposure to the VIX Index, one of the most common is the ZIV. Since these products take short positions on VIX Futures, the "contango effect" is a benefit, generating on average positive returns (see, Figure 1 and 2).

The objective of this paper is to test and exploit the phenomenon known as "volatility premium" using different volatility vehicles that aim to obtain exposure to the underlying in different ways. The main innovation comes from the fact that we introduced a Hidden Markov model (HMM) to capture structural breaks in the dynamics of volatility, and thus, trying to improve the investment decisions. The structure of the model identifies two states of the nature, one in which the volatility spread is high and hence understood to be expensive, and another where it decreases and understood to be cheap. By implementing a HMM, we aim at anticipating these shifts in the volatility behavior to improve our decision rules, avoiding the high risks conveyed with this kind of strategies and at the same time, being able to exploit and capture the promissory returns that could be achieved.

A HMM constitutes a statistical model which works with a Markov process with hidden states. The transition between the different states is governed by the transition probabilities. Besides, the model generates emissions probabilities which are observable and depends on the hidden states.[10]

One of the most used methods to estimate a HMM is the Baum-Welch algorithm (BM), proposed by Baum and Welch (1970), using a maximum likelihood approach. This method constitutes a particular case of the expectation maximization algorithm (EM), used several years before in different fields of science (see, for example, [Chis and Harrison, 2015], [Dempster, Laird and Rubin, 1977], [Levinson, 2005]). A similar method that is usually used in practice is the Hamilton filter (HL), although several studies have shown certain advantages of the BM over the HL.[11]

As for the decision making, we used the Viterbi algorithm, proposed in 1967 by Andrew Viterbi, which constitutes a decoding methodology using a dynamic programming algorithm. This method is used to find the most probable sequence of hidden states given the HMM model parameters previously estimated and the sequence of observable states.

[10]For detailed explanations on this model, see, for example, [Roman, Mitra and Spagnolo, 2010], [Shi and Weigend, 1997], [Visser, 2011], [Zhu *et al*, 2017].

[11] For more literature, see, for example, [Hamilton, 2010], [Mitra and Date, 2010], [Psaradakisa and Solabcd, 1998].

Since we have not found articles focused on capturing the volatility premium and its dynamics using a HMM, at least from an investor's point of view, we recognize that this paper is a valuable contribution to the current literature. One of the most attractive findings behind volatility trading comes from the impressive returns that an investor could achieve. However, the risk associated with this kind of strategies can be huge if an appropriate risk management is missing. For instance, during low volatility environments, selling volatility strategies can achieve outstanding performance. However, under markets shifts, where the conditions change abruptly, losses could be devastating for an investment portfolio.

2 Data and methodology

This paper focuses on trading volatility through volatility-based ETNs. To calculate the realized volatility, we employed daily returns of the Standard & Poor's 500 Index. Besides, we used the CBOE VIX Index as a proxy of the implied volatility. For trading purposes, we employed the Daily Inverse VIX Medium-Term ETN (ZIV) and the iShares 7-10 Years Treasury bond ETF (IEF).[12]

The sample period spans from April 2004 to August 2017. However, given that the ZIV ETN is available to trade since 2010, we replicated it from 2004 onwards by reproducing its methodology for backtesting purposes. The HMM employed consists in a discrete sequence of states with an ergodic Markov chain, meaning that the model can switch to every state without constraints. Besides, we assume a homogeneous Markov chain. One of our goals is to study the dynamics of the difference between implied volatility and realized volatility, widely known as volatility gap or volatility spread:

$$Spread_t = \mu_{z_i}(t) + \epsilon_{z_i}(t) \tag{1}$$

Where $\mu_{z_i}(t)$ and $\varepsilon_{z_i}(t)$ are the mean and the error term at time t, respectively. Both are conditional to the current regime z_i. Besides,

$$\varepsilon_t \sim N(0, \sigma_z) \tag{2}$$

Following the calculated spread, we made our trading rules. Depending on the current state dictated by the calibrated hidden Markov model, we define the corresponding trading decision:[13]

[12] Data about ETF of american Treasuries and of the ETN ZIV come from Yahoo Finance. As for the ZIV, given there is data only from 2010 November, for the previous years we replicated the ETN data using continuous rolling of futures with information from Quandl services (inside Quandl, the data supplier is Wiki Continuous Futures).

[13] (in terms of rebalancing), becoming a conservative methodology. On average, costs were 15 basis point for each rebalancing and we assume an initial portfolio of USD 100K.

• take short positions in VIX futures (trough the ZIV ETN), acting as a volatility seller, or

• allocate the total money available in long-term Treasury Bills (trough the IEF ETF).

2.1 Hidden Markov Model

As we said above, the hidden state sequence follows a Markov chain which satisfies the Markov property:

$$P[Z(t_n) = j|Z(t_1)\ldots .Z(t_{n-1}) = i] = P\left[Z(t_n) = j \mid Z(t_{n-1}) = i\right] \tag{3}$$

Besides, there are non-hidden continuous observations which are state-dependents of the current hidden state, which also satisfy the Markov property. The first task is to set the initial model parameters $\lambda = (\pi, \Gamma, \beta)$, where π is the initial probability of the state i, β are the emission probabilities, and Γ are the state-transition probabilities. The initial state probability is merely the probability of being in the state i at the beginning of the period under analysis. It is defined as following:

$$\pi_i = P(z_1 = i) \tag{4}$$

The matrix of transition probabilities indicates the probability of being in the state i and hold in that state or switch to another state and vice versa.

$$\Gamma = \begin{bmatrix} \alpha_{i,i} & \alpha_{i,j} \\ \alpha_{j,j} & \alpha_{j,i} \end{bmatrix} \tag{5}$$

Where the elements on the first row are defined as[14]

$$\alpha_{i,i} = P\left(z_{t+1} = i |\, z_t = i\right) \tag{6}$$

$$\alpha_{i,j} = P\left(z_{t+1} = j |\, z_t = i\right) \tag{7}$$

Since each row contains the conditional probabilities of being in the same state and the likelihood of switching to another state, the sum of them must be one. So, for both rows, the following constraint must be satisfied:

$$\sum_{j=1}^{N} \alpha_{i,j} = 1 \tag{8}$$

Where N is the number of hidden states, i represent the current state and j represents the next state given the current state. Finally, the emission probabilities

[14] For the second row, is the opposite.

are the probabilities to see an observable state at a specific point in time given the current hidden state.

$$b_i(j) = P\left(x_t = j \middle| \, z_t = i\right) \tag{9}$$

The initialization of the parameters constitutes an important step, because depending on that the results could be very different. That is why several approaches have been proposed to do that. For the subsequent periods, these parameters have to be re-estimated (or updated) by maximizing the probability $P\left(X \middle| \, \lambda\right)$. To do that, we used the expectation maximization algorithm mentioned above, which is a standard tool used in other fields of science to estimate the model parameters by maximum-likelihood when the data is not complete (in our problem, it is equivalent to the hidden states). Finally, for decoding and inference the most probable state, i.e., find the most likely sequence of hidden states that produce a given series of observation states, we used the Viterbi algorithm.

2.2 Model estimation: Baum-Welch Algorithm

One of the most challenge problems of HMMs is to estimate the model parameters. To tackle this problem, we used the Baum-Welch algorithm (see, for example, [Baum *et al*, 1970], [Matsuyama, 2003], [Matsuyama, 2011]), a particular case of the expectation-maximization algorithm. This method utilizes an iterative procedure which train the model parameters by maximizing the log-likelihood by an iterative procedure.

It is important to keep in mind that this method depends to a large extent on the initial predefined parameters. Besides, since the maximization problem is non-convex, the model does not achieve a global maximum, although guarantee reaches a local maximum. Given the initial parameters of the model and a sequence of observations, the algorithm re-estimates the λ parameters by maximizing the likelihood of

$$\lambda^* = P\left(X \mid \lambda\right) \tag{10}$$

We start defining the probability of the state i at time t and the probability of the state j at time $(t+1)$ given the observation sequence X and the initialized model parameters λ as following:

$$\phi_t\left(i,j\right) = P\left(z_t = i, \; z_{t+1} = j \mid X, \; \lambda\right) \tag{11}$$

$$= \frac{P\left(z_t = i, \; z_{t+1} = j \; , X \middle| \; \lambda\right)}{P\left(\; X \; \middle| \lambda\right)} \tag{12}$$

As the Baum-Welch method follows a forward-backward procedure, we define forward and backward variables. In one side, the forward variable is the following:

$$f_t\left(i\right) = P\left(x_1, x_2, x_3 \ldots ..x_t, z_t = i \mid \lambda\right) \tag{13}$$

Which is the joint probability of a determined observation sequence and a state given the model. As we can estimate the forward probability, we can know the objective function $P(X|\lambda)$ as the sum of the forward probability of each state. Following, by doing these calculations we can get the forward probability one step ahead:

$$f_{t+1}(i) = \left[\sum_{i=1}^{M} f_t(i)\ \alpha_{i,j} \right] b_i(X_{t+1}) \tag{14}$$

On the other side, the backward procedure gives us the following probability

$$\beta_t(i) = P(x_{t+1}, x_{t+2}, x_{t+3...}x_T \mid Z_t = i\ ,\ \lambda) \tag{15}$$
$$\beta_{t+1}(i) = P(x_{t+2}, x_{t+3}, x_{t+4...}x_T \mid Z_{t+1} = i\ ,\ \lambda) \tag{16}$$

Which is the probability of a set of the observed sequence of observations from $(t+1)$ to T given the model λ and the state Z_t at time t. Finally, we redefine $\phi_t(i,j)$ as following:

$$\phi_t(i,j) = \frac{f_t(i)\,\alpha_{i,j} b_j(X_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{R}\sum_{j=1}^{R} f_t(i)\,\alpha_{i,j} b_j(X_{t+1})\beta_{t+1}(j)} \tag{17}$$

Now, we define the component i of the transition probabilities matrix as following:

$$\Gamma_t(i) = p(z_t = i|X,\lambda) \tag{18}$$
$$\Gamma_t(i) = \sum_{i=1}^{R} \phi_t(i,j) \tag{19}$$

Summing $\Gamma_t(i)$ from $(t = 1)$ to $(T - 1)$ we obtain the expected number of state transition occurred from state i, defined as following:

$$\gamma\ (i) = \sum_{t=1}^{T-1}\sum_{i=1}^{R} \phi_t(i,j) \tag{20}$$

Besides, if we sum the above equation until T instead of $(T-1)$, we get the opposite. The expected number of state transition from another state to i, is defined as follows:

$$\theta\ (i) = \sum_{t=1}^{T}\sum_{i=1}^{R} \phi_t(i,j) \tag{21}$$

Now, we can estimate the model parameters $\lambda = (\pi, \Gamma, \beta)$. First, the expected numbers of times in state i is defined by π_i as

$$\pi_i = p(z_1 = i|X,\lambda) \tag{22}$$

The matrix of transition probabilities, which contains the probability of being in the state i and switch to the state j, is estimated by dividing the number of times

t that such transition has occurred by the expected number of transitions occurring from state i:

$$\alpha_{i,j} = \frac{\sum_t^{T-1} \phi_t(i,j)}{\gamma\ (i)} \tag{23}$$

Finally, regarding the emission probabilities, we can express $\bar{b}_t(x)$, as the probability of being in state j and observe a variable x divided by the expected times in state j

$$\bar{b}_t\ (x) = \frac{\sum_{t=1}^{T} \Gamma_t\ (j)'}{\theta\ (i)} \tag{24}$$

To summarize, the Baum-Welch algorithm starts with the initial values of the model λ. As a second step, such parameters are re-estimated on each iteration while $P(X|\lambda)$ is increasing. The model stops when that probability stops to increasing or when the defined maximum number of possible iterations is reached.

2.3 Inferring the most-likely state: The Viterbi algorithm

The Viterbi algorithm (VA) consists on a recursive method utilized to find the most probable sequence of hidden finite-states given the observation sequence and the estimated HMM λ. To do that, the algorithm search for the point of the function domain that maximize it using the argument of the maxima operator (arg max).

$$Z^* = argmax\ P(Z|X,\lambda) \tag{25}$$

$$argmax \prod_{t=0}^{T+1} P(X_t|Z_t)P(Z_t|Z_{t-1}) \tag{26}$$

Finally, based on the most probably state inferred by the above algorithm explained, we make the trading decisions.

3 Main Results

3.1 Offline approach

We started our analysis by estimating and running the model on an offline fashion, which is not useful for testing a trading strategy but instead give us insight about the behavior of the variable under analysis in different market conditions.

Besides, by studying the problem on this way, we find that the model was able to capture the changes in the volatility regimes. On one hand, during turbulent markets, generally driven by financial crises and political issues, among others, characterized by high volatility and huge losses in financial markets, the spread between the implied volatility and realized volatility tends to present high volatility and a negative mean, which imply a negative "volatility risk premium". On the other hand, when the markets are tranquil, characterized by a low volatility environment and high asset returns, the spread seems to be higher and low volatile. This regime presents positive "volatility risk premium".

Concerning trading, since our strategy gets returns by selling volatility, the regime with high spread and low volatility presents better conditions since high spread means that the investors are overpaying for assurance (relative to conventional pricing models), pushing up the implied volatility. Besides, considering the low volatility in the spread, the risk to suffer an abrupt increase in volatility is lower than in the other market states. We conducted this analysis using both, daily and weekly data. For the daily data, the matrix of transition probabilities estimated was the following:

$$\Gamma = \begin{bmatrix} 0.979 & 0.021 \\ 0.056 & 0.944 \end{bmatrix} \tag{27}$$

The response parameters or the observations equations in each state were the following:

$$\begin{cases} Spread_1 = 5.10 \; with \; \epsilon_1 \sim (0, 2.50) \\ Spread_2 = 1.88 \; with \; \epsilon_2 \sim (0, 10.28) \end{cases} \tag{28}$$

The duration of each state given that it is in the current state was calculated using the matrix of transition probabilities as follows:

$$E\left[Dur\right] = \frac{1}{1 - \alpha_{i,i}} \tag{29}$$

The expected duration of the state 1 is approximately 49 days while the expected duration of the state 2 is less than 18 days. Besides, the regime 1 is detected almost 75% of the time under analysis.

Regarding the weekly data, the results are quite similar. The matrix of transition probabilities obtained is the following:

$$\Gamma = \begin{bmatrix} 0.979 & 0.021 \\ 0.141 & 0.859 \end{bmatrix} \tag{30}$$

The response parameters or the observations equations in each state were the following:

$$\begin{cases} Spread_1 = 4.87 \; with \; \epsilon_1 \sim (0, 3.40) \\ Spread_2 = -0.82 \; with \; \epsilon_2 \sim (0, 12.57) \end{cases} \tag{31}$$

The expected duration of the state 1 is 48 weeks while the expected duration of the state 2 is more than 7 weeks. Besides, the regime 1 is detected almost 90% of the time under analysis. We were able to appreciate both volatilities spreads regimes, with a significant difference in the mean and the standard deviation. Besides, the average duration of each state seem to be reasonable, since calm markets, characterized by high volatility spread tend to be longer than distressed markets.

3.2 Online approach

To make the model useful for trading purposes, we follow an "online fashion", re-estimating the model with a lookback period of two years every day. As a decision rule, as we said above, we used the Viterbi algorithm to make inference about the most probable state.

Three different rules were imposed. First, we leave the strategy free, taking positions based on the daily inference about the most probable state. Secondly, we imposed the restriction that when a change of state was detected, the positions were not modified for at least one week. Finally, we restrict our strategy to monthly rebalances.

When the most probable state was a calm market, characterized by a high volatility spread, we take long positions in medium-term inverse volatility ETNs (ZIV), gaining negative exposure to the implied volatility. Opposite, during volatile markets, we allocate the total available money to 7-10 years Treasury Bills.

When we compare our strategy results against buy-and-hold the ZIV ETN and the S&P 500 index, we were able to notice an impressive improvement in the performance, with CAGRs of 19.26%, 20.30%, and 20.75% and annualized volatilities of 21.67%, 21.38%, and 19.24% for the daily, weekly, and monthly strategies, respectively, achieving Sharpe ratios near to 1 in all cases. Regarding the maximum drawdown, it was in all cases approximately -30%.

On table 2, we can appreciate the performance metrics of the proposed benchmarks. The CAGRs were 8.62% and 10.73% with volatilities of 18.89% and 30.53% for both S&P 500 and ZIV ETN, achieving Sharpe ratios of 0.43 and 0.35, respectively. Regarding the drawdowns were -54.63% and -86.43% for the S&P 500 and the ZIV ETN (see, Table 2)

Then, we divided the sample into two sub-periods. The first spans from April 2004 to January 2010, and the second from January 2010 to August 2017. In both sub-samples the proposed strategy outperformed all benchmarks. For the first, the returns achieved were 8.71%, 2.24%, and -4.88% with a volatility of 13.96%, 22.72%, and 27.92% for the strategy, the S&P 500 Index and the ZIV ETN passive investment. Regarding the maximum drawdowns, were -14.71%, -54.63%, and -86.43%, respectively. As for the second sub-sample, the strategy also outperformed the benchmarks (see, Table 3 and 4).

We also analyzed the strategy exposure over the whole sample to systematic risk factors, applying the Carhart four-factor model.

$$R_{i,t} = \alpha_i + \beta_{1,i} R_{1,t} + \cdots + \beta_{N,i} R_{N,t} + \epsilon_{i,t} \tag{32}$$

here $R_{i,t}$ is the excess returns on stock i at time t and N $\in$ [1, N] are the N factors; market, size, value and momentum.[15] The performance was also tested against the three-factor model of Fama and French, achieving similar results.

The strategies tested showed significant exposure to the outlined factors in statistical terms, except for the momentum factor. Additionally, we analyzed the alphas obtained by the strategies, achieving for all cases positive and significant alphas (statistic t greater than 2) (see, Table 5).

4 Recent events

In recent times, ETNs negatively exposed to volatility have proliferated because they offer a high rate of risk-adjusted returns. This led to a large increase and arrival of funds with focus and specialization in this asset class. The volatility trade seemed to be quite simple; just taking long positions in negatively exposure ETNs and letting profits run. However, as people say, "nothing lasts forever". Here was not the exception.

4.1 Facts

On February 5 there was a strong sell-off, which led to a substantial increase in the VIX index and, therefore, tons of losses to short sellers of volatility. The magnitude was so high that some negatively exposed ETNs reached a liquidation event and disappeared, e.g., the XIV ETN. However, when analyzing the event concerning absolute volatility, it could be argued that during the past, there were many events with such spikes in volatility (see, Figure 3). The same can be noticed when the spread between the VIX index and the realized volatility is analyzed (see, Figure 4).

However, when analyzing another critical variable such as the ratio between the realized volatility (30-day close-to-close in our case) over the previous level of the VIX Index, it is possible to appreciate that the Index, considered a forward-looking

[15] For more literature on multi-factor models, see, for example, [Carhart, 1997], [Fama and French, 1993], [Fama and French, 2012].

measure of the volatility, does not expect this event. In figure 6 we can appreciate that the ratio suffers an abruptly change, reaching one of the highest peaks from 2000, similar to the levels reached in the Great Financial Crisis and other two "volatility events" (2011 and 2015). Based on these findings, one could argue that this peak in volatility was different for other events in the sense that the market was not expecting it. This phenomenon was studied by Cliff Asness, who denominated this as "surprise factor" (see, [Asness, 2018]).

Also, we also should analyze the futures market. Figure 5 shows the abrupt change in the market risk perception through the VIX futures curve, changing from contango to backwardation in just a few days. This is another essential variable in considering when studying volatility markets and changes in risk perceptions.

4.2 How did the proposed model perform in such events?

As a final thought, we noticed that our proposed model was able to identify this abruptly break in the behavior of the spread, which in turns generates a switch signal from short VIX to long Treasuries (see Figure 7). However, we believe that by including other variables, such as these stated in the Section 4.1 it is possible to improve the performance along the time.

Conclusions

As we introduced in the abstract, selling insurance can become a profitable strategy if an appropriate risk management methodology is under place. The data confirms that the VIX Index is generally above the 30-day rolling volatility giving rise to the volatility premium. We have developed a regime switching approach by using a Hidden Markov Model with two states of nature, one in which volatility is understood to be expensive, so it becomes profitable to sell it, and another one in which it becomes cheap, so it is better to hold positions in U.S. Treasuries. Based on the inferences obtained with the proposed model, we set a trading strategy aimed to profit from the evidence found and set an algorithm that sells volatility when expensive, and switches to U.S. Treasuries when cheap. The strategy is facilitated by the variety of financial contracts available to trade volatility. The results show that excess returns regarding the Carhart four-factor model can be obtained.

As we also mentioned in the review, the volatility premium does exist and can be associated with insurance against events that have never happened before and may occur in the future, or a market microstructure biased towards buyers. Even though we were not able to explain, we have developed a different mechanism to measure it and take trading advantage. For future work stands a chance to continue researching the causes behind the spread, and to analyze the performance of innovative financial contracts to arrive at further conclusions. Uncertainty is a matter that becomes difficult to grab and resorts to emotions such as greed and fear.

Further research

In this research, we have shown that by trading volatility-related products, it is possible to obtain excessive returns, measured with traditional risks factors models such as the Fama French model of three factors (not reported) and the Carhart-four factor model by employing the appropriate risk management tools. E.g., by using a Hidden Markov model aimed at modeling the dynamics of the spread between implied volatility and realized volatility, we have shown that different states with structural break can be identified and exploited from a trading perspective. However, we believe that there is an ample space to improve, mainly on two sides; on the one hand, adding more features (or predictive) variables to the model and not limiting to the "volatility spread" as proposed here. In this regard, the slope and the curvature of the VIX futures curve are just two examples of the variables that we believe that could help to improve the performance. On the other hand, we see an ample space for improvement regarding the model, where there is extensive literature with different modeling approaches which are use in other fields of the science, e.g., techniques borrowed from machine learning.

References

Andersen, T.G., Fusari, N. and Todorov, V. (2015) The risk premia embedded in index options. *Journal of Financial Economics* 117(3): 558-584.

Bakshi, G. and Kapadia, N. (2003) Volatility Risk Premiums Embedded in Individual Equity Options. *Journal of Derivatives* 11(1): 45-54.

Baum, L.E., Petrie, T., Soules, G. and Weiss, N. (1970) A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *Annals of Mathematical Statistics* 41(1): 164-171.

Carhart, M.M. (1997) On Persistence in Mutual Fund Performance. *Journal of Finance* 52(1): 57-82.

Chis, T. and Harrison, P.G. (2015) Adapting Hidden Markov Models for Online Learning. *Electronic Notes in Theoretical Computer Science* 318: 109-127.

Asness, A. (2018) Wild but Not Crazy. *AQR Working Paper.*

Dapena, J.P. and Siri, J.R. (2015) Index Options Realized Returns distributions from Passive Investment Strategies. *UCEMA Working Paper.* Available online: `https://ssrn.com/abstract=2733774`.

Dapena, J.P. and Siri, J.R. (2016) Testing Excess Returns from Passive Options Investment Strategies. *UCEMA Working Paper.* Available online: `https://ssrn.com/abstract=2998787`.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society* 39(1): 1-38.

Eraker, B. (2007) The Volatility Premium. *Working Paper.*

Fama, E.F. and French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1): 3-56.

Fama, E.F. and French, K.R. (2012) Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105(3): 457-472.

Fleming, J. (1999) The economic significance of the forecast bias of S&P 100 index option implied volatility. *Advances in Futures and Options Research* 10: 219-251.

Hamilton, J.D. (2010) Regime switching models. In: Durlauf S.N. and Blume L.E. (eds.) *Macroeconometrics and Time Series Analysis.* (5th ed.). London: Palgrave Macmillan.

Harvey, C.R. and Siddique, A. (2000) Conditional Skewness in Asset Pricing Tests. *Journal of Finance* 55(3): 1263-1295.

Ilmanen, A. (2012) Do Financial Markets Reward Buying or Selling Insurance and Lottery Tickets? *Financial Analysts Journal* 68(5): 26-36.

Ilmanen, A. (2013) "Do Financial Markets Reward Buying or Selling Insurance and Lottery Tickets?": Author Response. *Financial Analysts Journal* 69(2): 19-21.

Jackwerth, J.C. and Rubinstein, M. (1996) Recovering Probability Distributions from Option Prices. *Journal of Finance* 51(5): 1611-1631.

Kozhan, R., Neuberger, A. and Schneider, P. (2013) The Skew Risk Premium in the Equity Index Market. *Review of Financial Studies* 26(9): 2174-2203.

Kraus, A. and Litzenberger, R.H. (1976) Skewness Preference and the Valuation of Risk Assets. *Journal of Finance* 31(4): 1085-1100.

Levinson, S.E. (2005) Mathematical Models for Speech Technology. Chichester, UK: John Wiley & Sons, Ltd.

Lin, B.-H. and Chen, Y-.J. (2009) Negative Market Volatility Risk Premium: Evidence from the LIFFE Equity Index Options. *Asia-Pacific Journal of Financial Studies* 38(5): 773-800.

Matsuyama, Y. (2003) The α-EM algorithm: Surrogate likelihood maximization using α-logarithmic information measures. *IEEE Transactions on Information Theory* 49(3): 692-706.

Matsuyama, Y. (2011) Hidden Markov model estimation based on alpha-EM algorithm: Discrete and continuous alpha-HMMs. In: *International Joint Conference on Neural Networks.*, pp. 808-816.

Mitra, S. and Date, P. (2010) Regime switching volatility calibration by the Baum-Welch method. *Journal of Computational and Applied Mathematics* 234(12): 3243-3260.

Psaradakisa, Z. and Solabcd, M. (1998) Finite-sample properties of the maximum likelihood estimator in autoregressive models with Markov switching. *Journal of Econometrics* 86(2): 369-386.

Roman, D., Mitra, G. and Spagnolo, N. (2010) Hidden Markov models for financial optimization problems. *IMA Journal of Management Mathematics* 21(2): 111-129.

Shi, S. and Weigend, A.S. (1997) Taking time seriously: hidden Markov experts applied to financial engineering. In: *Proceedings of the IEEE/IAFE Conference on Computational Intelligence for Financial Engineering.*

Sinclair, E. (2013) Volatility Trading. Hoboken, NJ: John Wiley & Sons, Inc.

Taleb, N.N. (2013) "Do Financial Markets Reward Buying or Selling Insurance and Lottery Tickets?": A Comment. *Financial Analysts Journal* 69(2): 17-19.

Visser, I. (2011) Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *Journal of Mathematical Psychology* 55(6): 403-415.

Zhu, D.-M., Ching, W.-K., Elliot, R.J., Siu, T.-K. and Zhang, L. (2017) Hidden Markov models with threshold effects and their applications to oil price forecasting. *Journal of Industrial & Management Optimization* 13(2): 757-773.

Graphs and tables

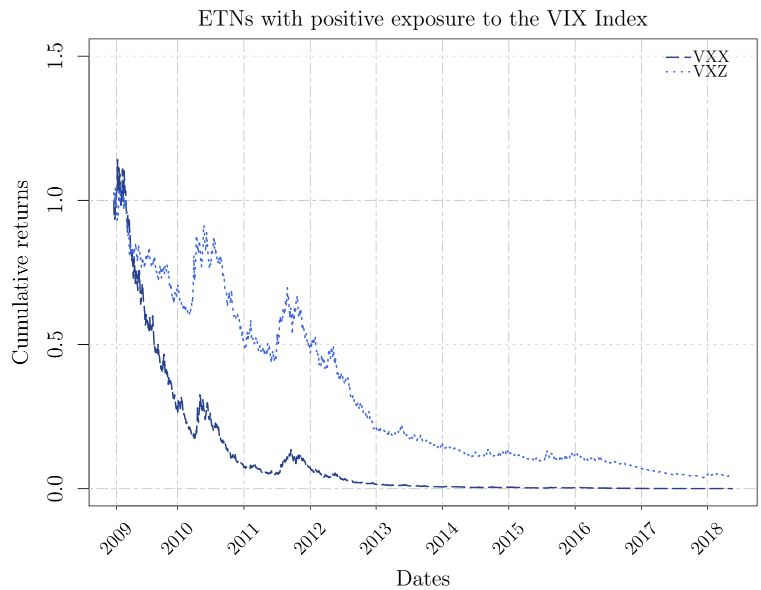

Figure 1: This plot presents ETNs with positive exposure to the VIX Index. It is possible to appreciate how both series have been losing value over time due to the “roll” or “contango” costs.

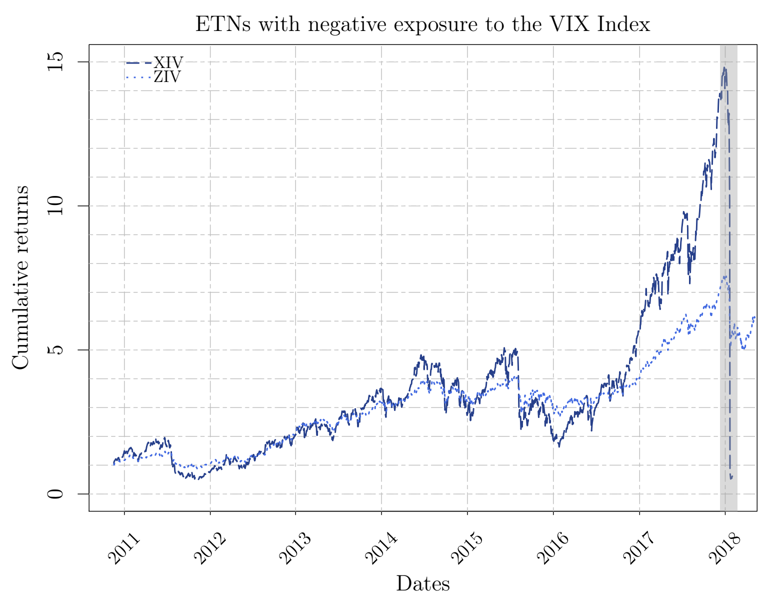

Figure 2: This plot presents ETNs with negative exposure to the VIX Index. Unlike the VXX and VXZ ETNs, the volatility-products with negative exposure to it do not suffer the contango costs.

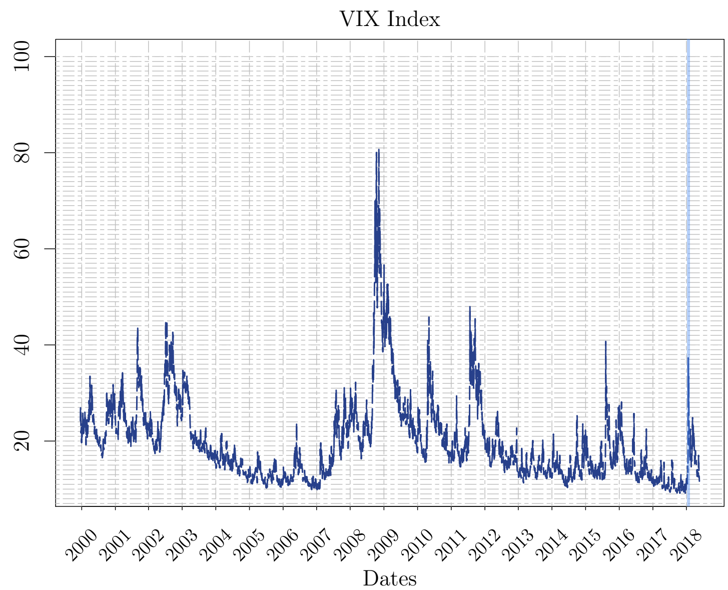

Figure 3: This plot presents the VIX Index from January 2000 to June 2018. The shaded area represents the "volatility event" under study on February 5.

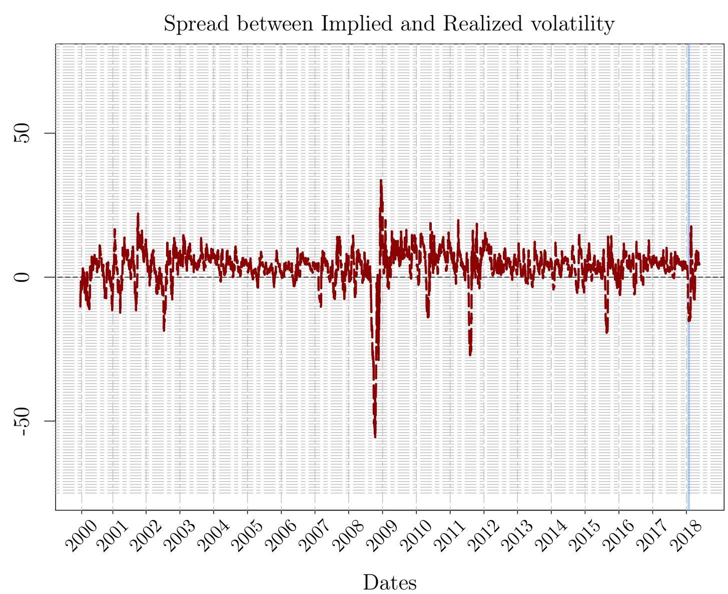

Figure 4: This plot presents the spread between the previous-month level of the VIX Index and the rolling 30-day close-to-close volatility from January 2000 to June 2018. The shaded area represents the "volatility event" under study on February 5.

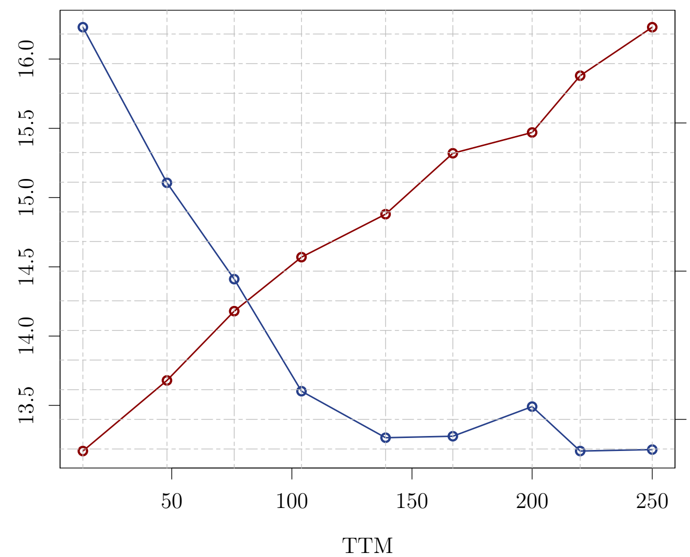

Figure 5: This plot presents the VIX futures curve before (red) and after (blue) the "volatility event" under study.

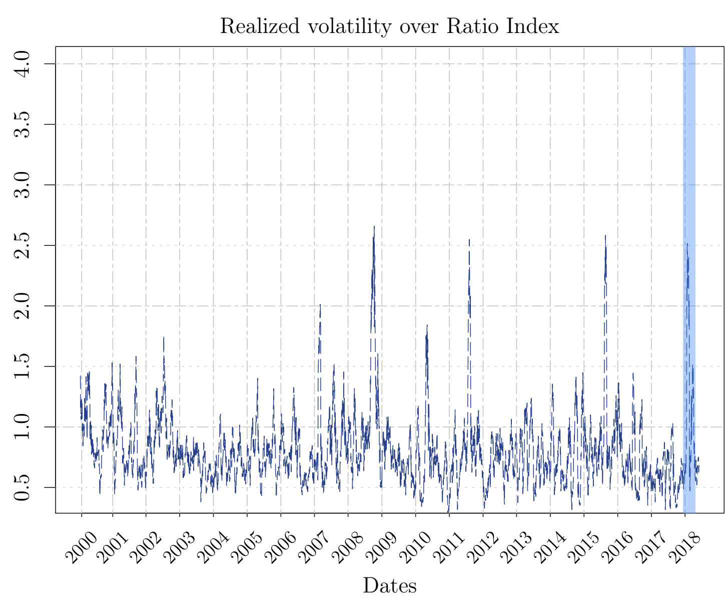

Figure 6: This plot presents the ratio between the rolling 30-day close-to-close realized volatility over the previous-month level of the VIX Index. It is possible to appreciate the "surprise factor" during the particular event at the beginning of February 2018.

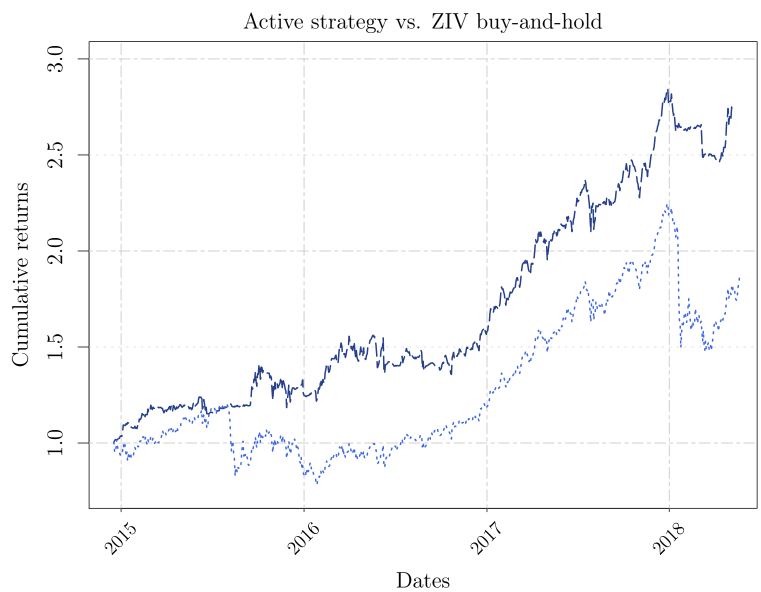

Figure 7: This plot presents the equity curve of the strategy from 2015 to 2018.

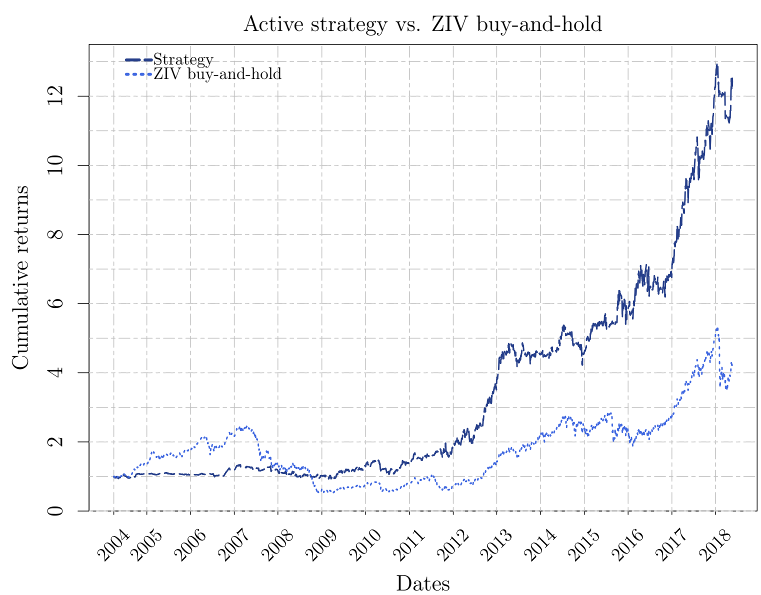

Figure 8: This plot presents the equity curve of the strategy during the whole period.

	Daily	Weekly	Monthly
CAGR	19.26%	20.30%	20.75%
SD	21.67%	21.38%	19.24%
Sharpe ratio	0.89	0.95	1.08
Skewness	-0.56	-0.55	-0.55
Kurtosis	7.39	7.82	9.69
MaxDD	-35.36%	-34.43%	-29.50%
Calmar Ratio	0.53	0.59	0.70

Table 1: shows the descriptive statistics and ratios on annual basis of the proposed strategies under different rebalance periods (daily, weekly, and monthly) for the whole sample (from 2004/04/01 to 2018/06/04).

	S&P 500	ZIV ETN
CAGR	8.62%	10.73%
SD	19.89%	30.53%
Sharpe ratio	0.43	0.35
Skewness	0.19	-0.57
Kurtosis	16.91	4.02
MaxDD	-54.63%	-86.43%
Calmar Ratio	0.11	0.07

Table 2: shows the descriptive statistics and ratios on annual basis of the proposed benchmarks for the whole sample (from 2004/04/01 to 2018/06/04).

	Strategy	S&P 500	ZIV ETN
CAGR	8.71%	2.24%	-4.88%
SD	13.96%	22.72%	27.92%
Sharpe ratio	0.62	0.11	-0.17
Skewness	-0.49	0.42	-0.65
Kurtosis	12.17	16.35	5.04
MaxDD	-14.71%	-69.86%	-86.43%
Calmar Ratio	0.59	0.03	-0.04

Table 3: shows the descriptive statistics and ratios on annual basis of the proposed strategy and benchmarks for the first sub-sample sample (from 2004/04/01 to 2010/01/01).

	Strategy	S&P 500	ZIV ETN
CAGR	30.26%	13.34%	23.30%
SD	24.03%	14.80%	30.58%
Sharpe ratio	1.25	0.76	0.90
Skewness	-0.51	-0.38	-0.53
Kurtosis	7.25	4.30	3.41
MaxDD	-34.51%	-19.00%	-51.12%
Calmar Ratio	0.86	0.70	0.45

Table 4: shows the descriptive statistics and ratios on annual basis of the proposed strategy and benchmarks for the second sub-sample sample (from 2010/01/01 to 2018/06/04).

	Alpha	RMRF	SMB	HML	MOM	Adj. R^2
Daily	0.0006	0.451	0.258	0.1147	0.031	0.191
	(2.93)	**(22.39)**	**(6.68)**	**(2.82)**	(1.15)	
Weekly	0.0006	0.425	0.257	0.1168	0.041	0.176
	(2.83)	**(21.23)**	**(6.72)**	**(2.89)**	(1.54)	
Monthly	0.0006	0.289	0.209	0.098	0.035	0.107
	(3.22)	**(15.47)**	**(5.84)**	**(2.61)**	(1.43)	

Table 5: shows the results of regressing the strategies returns against the four factor-model under different rebalance periods (daily, weekly, and monthly) for the whole sample (from 2004/04/01 to 2018/06/04). Bolded t-stats are significant at 95% level.