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

Última versión: Noviembre 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

Electronic copy available at: https://ssrn.com/abstract=3257073

Electronic copy available at: https://ssrn.com/abstract=3257073

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

November, 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. In concordance with existing literature, we show evidence that VIX is generally above the 30-day rolling volatility giving rise to the volatility premium, so that selling volatility can become a profitable trading strategy as long as proper risk management is under place. As a contribution, we introduce the implementation of a Hidden Markov Model (HMM), identifying two and three states of the nature and showing that the volatility premium undergoes temporal breaks in its behavior. Based on this, we formulate different trading strategies by selling volatility and switching to medium-term U.S. Treasury Bills when appropriated. We test the performance of the strategies using the conventional Carhart four-factor model showing positive and statistically significant alphas, even after considering transaction costs.

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

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

[2] Juan A. Serur, MS in Finance, is an Assistant Professor at University of CEMA and the Academic Secretary of the Master in Finance. Email: juanandresserur@gmail.com

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

Electronic copy available at: https://ssrn.com/abstract=3257073

Introduction

Selling insurance can be 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 regard, models in finance have historically aimed at capturing and appropriately measuring risk and evaluating the proper expected returns that compensates for it. Investors can act as insurance providers by selling risk when they consider it is overpriced, i.e., expected returns do not compensate for risk, and buy it when it is supposedly underpriced, i.e., expected returns overcompensate for risk. Hence a model capable of identifying when risk is expensive can be very valuable. In most of the cases, risk is measured 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 models intent to eliminate the drift from the calculation to put a price on volatility. Due to the importance of the volatility in financial markets, 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 that 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 VIX is generally above the 30-day rolling realized volatility. This results in what is known as "volatility premium"[6].

The purpose of the present work is to formally test the volatility premium through a quantitative model with the intention to exploit the fact that sometimes volatility is overpriced (e.g., implied volatility is greater than realized volatility), creating opportunities to make a profit by selling it. We show how to exploit this phenomenon by using different volatility vehicles that intents to gain exposure to the underlying in different ways through a Hidden Markov Model (HMM) where two and three states of the nature are identified giving rise to trading opportunities.

[4] However, more advanced models of finance aim to capture superior moments of the distribution (see, for example, [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].

Electronic copy available at: https://ssrn.com/abstract=3257073

The paper proceeds as follows: 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 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 high. In that regard, some authors have shown that selling protection against the skewness can be profitable (see, for example, [Ilmanen, 2012], [Ilmanen, 2013]). An investor can systematically sell 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 of the VIX Index by the CBOE in 1993. The evidence, as mentioned in the previous section, shows that there is a spread between expected volatility and realized volatility, leading to the hypothesis 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 investors 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, being common to see that realized volatility is in some cases 30% lower than the implied volatility.[8] He quotes a number of reasons, such as 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 patterns in the data do not stop there. VIX futures, for example, tend to overestimate the future value of the spot value of the VIX Index.

This type of events leads to the "volatility risk premium". Since investors seek for protection and are reluctant to the exposure to downside tail risk, acting as an

[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.

Electronic copy available at: https://ssrn.com/abstract=3257073

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 explanatory power on the positive difference between implied and realized volatility (see, for example, [Jackwerth and Rubinstein, 1996]). Others showed than on average, Black-Scholes implied volatilities of close to at-the-money equity options tend to be higher than historical realized volatility (see, for example, [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 former and 16% the latter. This difference translates into excess returns for index option sellers (see, for example, [Eraker, 2007]).[9]

As mentioned above, many explanations about the volatility risk premium have been proposed. 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 other 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 through volatility-related products like ETNs.

There are mainly two groups of volatility-related ETNs; there are products with both positive and negative exposure to the underlying. 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.[10]

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", meaning that long-term futures prices are higher than short-term futures, implying a roll cost due to the daily rebalance that is translated into daily losses.

As for products with negative exposure to the VIX Index, one of the most com-

[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].

[10] For more information on the statistics of volatility products and their applications for trading, see [Avellaneda and Papanicolaou, 2018].

Electronic copy available at: https://ssrn.com/abstract=3257073

mon is the ZIV. Since these products take short positions on VIX Futures, the "contango effect" is beneficial, 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 get exposure to the underlying in different ways. The main innovation comes from the fact that we introduced a HMM to capture structural breaks in the dynamics of volatility, which contributes to improve the investment decisions. We calibrated two models, with two and three states of nature. In both cases we identified that the variable under analysis is state-dependent. By implementing a HMM, we aim to anticipate these shifts in the volatility behavior to improve our decision rules, avoiding the high risks conveyed with this kind of strategies.

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 dependent on the hidden states.[11]

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), which has been widely used 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.[12]

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.

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 suggest 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 can achieve. However, the risk associated with this kind of strategies can be huge if an appropriate risk management is missing. For instance, during low

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

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

Electronic copy available at: https://ssrn.com/abstract=3257073

volatility environments, selling volatility strategies can achieve outstanding performance. However, under markets shifts, where the conditions change abruptly, losses can be devastating for an investment portfolio.

2 Data and methodology

This paper focuses on trading volatility through volatility-based ETNs and ETFs. To calculate the realized volatility, we employed daily returns of the Standard & Poor's 500 Index. In addition, we used the CBOE VIX Index as a proxy of the implied volatility.

For trading purposes, we chose as investable alternatives two ETFs and one ETN. To be precise, VelocityShares Daily Inverse VIX Medium-Term ETN (Ticker: ZIV), ProShares Short VIX Short-Term Futures ETF (Ticker: SVXY) and iShares 7-10 Years Treasury Bonds ETF (Ticker: IEF). The sample period spans from April 2004 to October 2018. However, since ZIV and SVXY inception dates are after 2010, we replicated previous values with VIX futures, allowing us to extend the backtest up to 2004.[13]

In all cases, we used closing prices. In this regard, in order to avoid a *look ahead bias*, we set up all the positions during each rebalancing date one day after the signal was generated, that is, each signal is generated at time t and the transaction is executed at time $t+1$. As for the trading periodicity, the strategy of 2 states changes on average 16 times a year, that is, it rebalances the portfolio every 22 days. As for the strategy of the 3 states, it changes 55 times per year on average, which means a rebalancing approximately every 7 days. Considering that only 2 or 3 ETFs are traded for both strategies, respectively, transaction costs are ephemeral.[14]

As we mentioned above, one of our goals was 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,

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

[13] Data of ETFs and ETNs was extracted from Yahoo Finance. Data used for replication of these instruments was extracted from Quandl.

[14] On average, transaction costs were 10 basis point for each rebalancing and we assume an initial portfolio of USD 100K. This shows that this strategy is not only for large investors, but also for retail investors.

Electronic copy available at: https://ssrn.com/abstract=3257073

Following the calculated spread, we set our trading rules. Depending on the current state dictated by the calibrated HMM, we defined the corresponding trading decisions:

- take short positions in VIX futures (trough the ZIV ETN or the SVXY ETF), 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

In this section, we explain the development of the HMM for the case of 2 states. Regarding the case of 3 states, the model can be extended quite easily. The HMM employed consists of a discrete sequence of states with an ergodic Markov chain, meaning that the model can switch to every state without constraints. Besides, a homogeneous Markov chain is assumed, 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 dependents on the current hidden state, which also satisfies 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 remain 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[15]

$$\alpha_{i,i} = P\left(z_{t+1} = i \middle| z_t = i\right) \tag{6}$$
$$\alpha_{i,j} = P\left(z_{t+1} = j \middle| z_t = i\right) \tag{7}$$

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. Therefore,

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

Electronic copy available at: https://ssrn.com/abstract=3257073

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 represents the current state and j represents the next state given the current state. Finally, the emission probabilities are the probabilities of seeing an observable state at a specific point in time given the current hidden state.

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

The initialization of the parameters constitutes an important step, because depending on that the results could vary substantially. 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 $P\left(X \right| \lambda)$. 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 inferring the most probable state, i.e., find the most likely sequence of hidden states that produces a given series of observation states, we used the Viterbi algorithm.

2.2 Model estimation: Baum-Welch Algorithm

One of the most challenging 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 trains the model parameters by maximizing the log-likelihood by an iterative process.

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 guarantees 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:

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

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

Electronic copy available at: https://ssrn.com/abstract=3257073

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

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

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

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

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

$$\beta_t\left(i\right) = P\left(x_{t+1}, x_{t+2}, x_{t+3\ldots} x_T \mid\, Z_t = i\;,\;\lambda\right) \tag{15}$$

$$\beta_{t+1}\left(i\right) = P\left(x_{t+2}, x_{t+3}, x_{t+4\ldots} x_T \mid\, Z_{t+1} = i\;,\;\lambda\right) \tag{16}$$

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 $\phi_t\left(i,j\right)$ as following:

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

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

$$\Gamma_t\left(i\right) = p\left(z_t = i|X,\lambda\right) \tag{18}$$

$$\Gamma_t\left(i\right) = \sum_{i=1}^{R} \phi_t(i,j) \tag{19}$$

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

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

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

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

Electronic copy available at: https://ssrn.com/abstract=3257073

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 ceases to increase 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 of a recursive method utilized to find the most probable sequence of hidden finite-states given the sequence of observations and the estimated HMM λ. To do that, the algorithm searches for the point of the function domain that maximizes 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 probable state inferred by the algorithm explained above, 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 gave us insight about

Electronic copy available at: https://ssrn.com/abstract=3257073

the behavior of the variable under analysis in different market conditions.

Besides, by studying the problem in this way, we found 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 implies 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 less volatile. This regime presents a 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 insurance (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:

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

As for the 3 state model, the transition matrix was the following:

$$\Gamma = \begin{bmatrix} 0.940 & 0.038 & 0.021 \\ 0.016 & 0.946 & 0.038 \\ 0.002 & 0.053 & 0.946 \end{bmatrix} \tag{28}$$

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

$$\begin{cases} Spread_1 = 5.10 \ with \ \epsilon_1 \sim (0, 2.50) \\ Spread_2 = 1.88 \ with \ \epsilon_2 \sim (0, 10.28) \end{cases} \tag{29}$$

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

$$E\left[Dur\right] = \frac{1}{1 - \alpha_{i,i}} \tag{30}$$

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

Electronic copy available at: https://ssrn.com/abstract=3257073

75% of the time under analysis. Then, we tested a 3 stated model. The response parameters or the observations equations in each state were the following:

$$\begin{cases} Spread_1 = 7.95 \ with \ \epsilon_1 \sim (0, 2.59) \\ Spread_2 = 2.71 \ with \ \epsilon_2 \sim (0, 1.87) \\ Spread_3 = -6.96 \ with \ \epsilon_3 \sim (0, 12.91) \end{cases} \tag{31}$$

The expected duration of the state 3 is approximately 18 days while the expected duration of the state 2 is less than 19 days and the state 3 less than 15 days. Besides, the regime 3 is detected almost 40% of the time under analysis and the state 2 is 47%.

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

$$\Gamma = \begin{bmatrix} 0.979 & 0.021 \\ 0.141 & 0.859 \end{bmatrix} \tag{32}$$

As for the 3 state model, the transition matrix is the following:

$$\Gamma = \begin{bmatrix} 0.850 & 0.145 & 0.004 \\ 0.069 & 0.906 & 0.025 \\ 0.138 & 0.001 & 0.861 \end{bmatrix} \tag{33}$$

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

$$\begin{cases} Spread_1 = 4.87 \ with \ \epsilon_1 \sim (0, 3.40) \\ Spread_2 = -0.82 \ with \ \epsilon_2 \sim (0, 12.57) \end{cases} \tag{34}$$

The expected duration of the state 1 is 48 weeks while the expected duration of the state 2 is more than 7 weeks. In addition, the regime 1 is detected almost 90% of the time under analysis. As for the 3 state model, the response parameters or the observations equations in each state were the following:

$$\begin{cases} Spread_1 = 7.73 \ with \ \epsilon_1 \sim (0, 2.88) \\ Spread_2 = 2.59 \ with \ \epsilon_2 \sim (0, 2.70) \\ Spread_3 = -6.73 \ with \ \epsilon_3 \sim (0, 15.06) \end{cases} \tag{35}$$

The expected duration of the state 3 is 6 weeks while the expected duration of the state 2 is more than 11 weeks and the expected duration of the state 1 is 8 weeks. Besides, the regime 3 is detected almost 33% of the time under analysis and the regime 2 is 57% of the time. With this analysis, we were able to appreciate the different regimes, with a significant difference in the mean and the standard deviation. Besides, the average duration of each state seems to be reasonable, since calm markets, characterized by high volatility spread tend to be longer than distressed markets.

Electronic copy available at: https://ssrn.com/abstract=3257073

3.2 Online approach

To make the model useful for trading purposes, we followed an "online fashion", reestimating 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.

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

On table 1, we can appreciate the performance metrics of the proposed benchmarks during the whole sample. The CAGRs were 8.48%, 10.07% and 14.59% with volatilities of 18.28%, 29.48% and 61.66% for the S&P 500, the ZIV and the SVXY, with Sharpe ratios of 0.46, 0.34 and 0.24, respectively (see, Table 1).

When we compare our strategies results against the buy-and-hold strategies exposed above, we were able to notice an impressive outperformance, with CAGRs of 17.24%, 45.12%, and 28.78% and annualized volatilities of 21.74%, 43.70%, and 33.20% for the 2-states (ZIV), 2-states (SVXY), and 3-states strategies, respectively, achieving Sharpe ratios higher than 0.80 in all cases (see, Table 4).

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 October 2018. In both sub-samples the proposed strategy outperformed all benchmarks (see, Tables 2, 3, 5 and 6).

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

$$R_{i,t} = \alpha_i + \beta_{1,i} R_{1,t} + \cdots + \beta_{N,i} R_{N,t} + \epsilon_{i,t} \tag{36}$$

Here $R_{i,t}$ is the excess returns of the strategy i at time t and $N \in [1, \mathrm{N}]$ are the N systematic risk factors; market, size, value and momentum.[16] The performance was also tested against the three-factor model of Fama and French, achieving similar results (not reported).

The strategies tested showed significant exposure to almost all the outlined factors in statistical terms. Additionally, we analyzed the alphas obtained by the strategies, achieving for all cases positive and significant alphas (see, Table 7).

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

Electronic copy available at: https://ssrn.com/abstract=3257073

4 Recent events

On February 5 there was a strong sell-off, which led to a substantial increase in the VIX Index, generating substantial 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 we analyze the index historically, we can observe that 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 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, for example, [Asness, 2018]).

In addition, we 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 to consider when studying volatility markets and changes in risk perceptions.

4.1 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 for the two cases (2 and 3 states), which in turns generates a switch signal (see, Figures 7 and 8 to appreciate the equity curves). At the same time, we believe that by including other variables, such as those indicated in Section 4, there is room to further improve the performance of the strategies.

Electronic copy available at: https://ssrn.com/abstract=3257073

Conclusions

As we introduced in the abstract, selling insurance can be a profitable strategy if an appropriate risk management methodology is under place. 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 and three states of the nature. Based on the inferences obtained with model, we established different trading strategies aimed at obtaining benefits from the evidence found, i.e., selling volatility when it is perceived to be expensive and switching to U.S. Treasury bonds when it is cheap. The results show statistically significant alphas when the returns of the strategies are tested against the Carhart four-factor model.

As we mentioned in the review, the volatility premium can be associated with different causes, such as insurance against events that have never occurred and may occur in the future, market microstructure issues, among others. Although we could not explain the pure causes behind the volatility risk premium, and it is possible that this will never be fully explained, we have developed a different mechanism to measure it and take advantage through different trading strategies. As future work, there is the possibility to deepen the research of the causes behind the volatility spread and analyze the performance of innovative financial contracts to reach additional conclusions.

Further research

In this research, we have shown that by trading volatility-related products, it is possible to obtain excess 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. 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 we 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 used in other fields of science, such as, for example, techniques borrowed from machine learning.

Electronic copy available at: https://ssrn.com/abstract=3257073

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.

Avellaneda, M. and Papanicolaou, A. (2018) Statistics of VIX Futures and Their Applications to Trading Volatility Exchange-Traded Products. *Journal of Investment Strategies* 7(2): 1-32.

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.

Electronic copy available at: https://ssrn.com/abstract=3257073

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.

Electronic copy available at: https://ssrn.com/abstract=3257073

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.

Electronic copy available at: https://ssrn.com/abstract=3257073

Graphs and tables

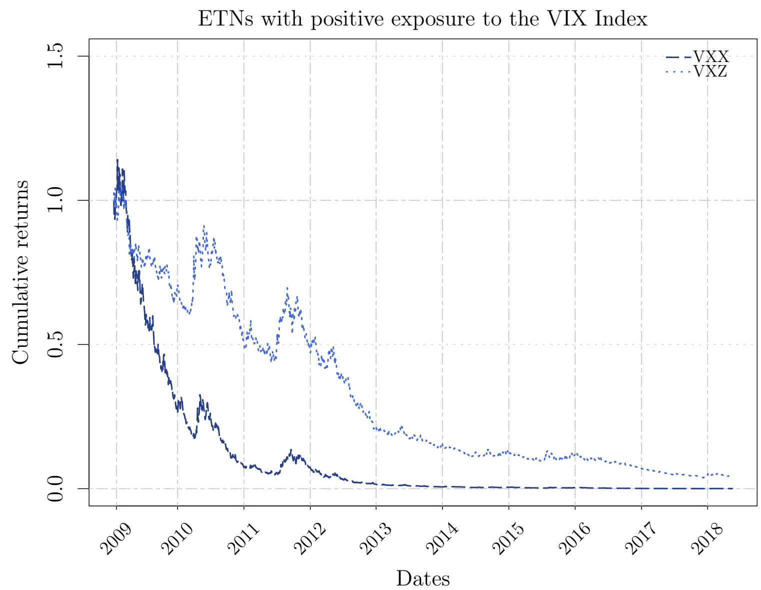

Figure 1: This plot presents ETNs with positive exposure to the VIX Index. It is possible to appreciate how both series have been losing value over time due to the “roll” or “contango” costs.

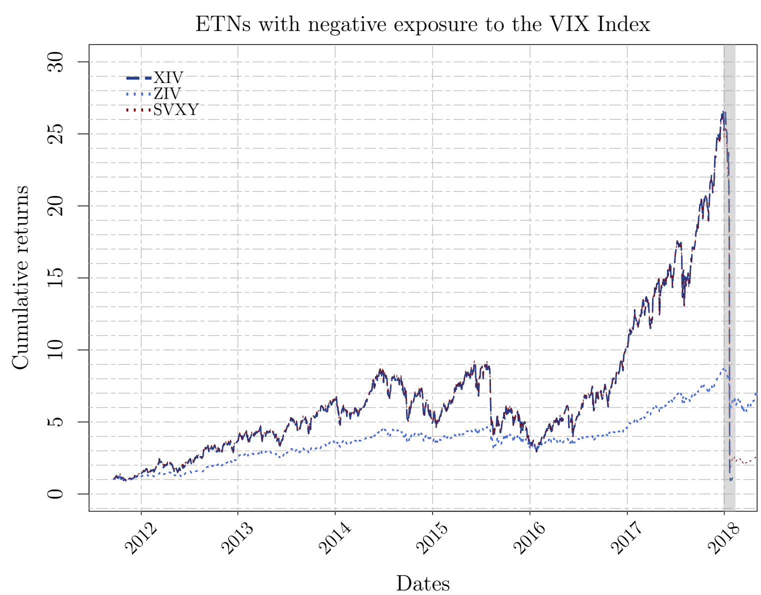

Figure 2: This plot presents ETNs with negative exposure to the VIX Index. Unlike the VXX and VXZ ETNs, the volatility-products with negative exposure to it do not suffer the contango costs.

Electronic copy available at: https://ssrn.com/abstract=3257073

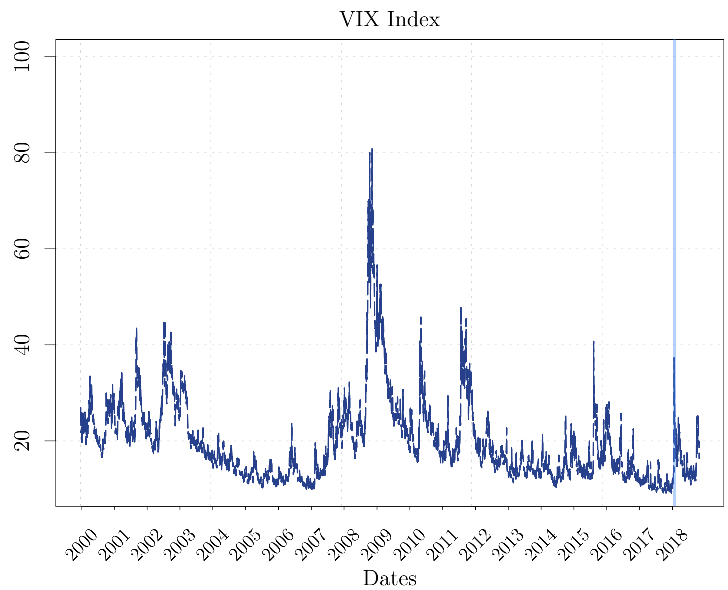

Figure 3: This plot presents the VIX Index from January 2000 to October 2018. The shaded area represents the “volatility event” under study on February 5.

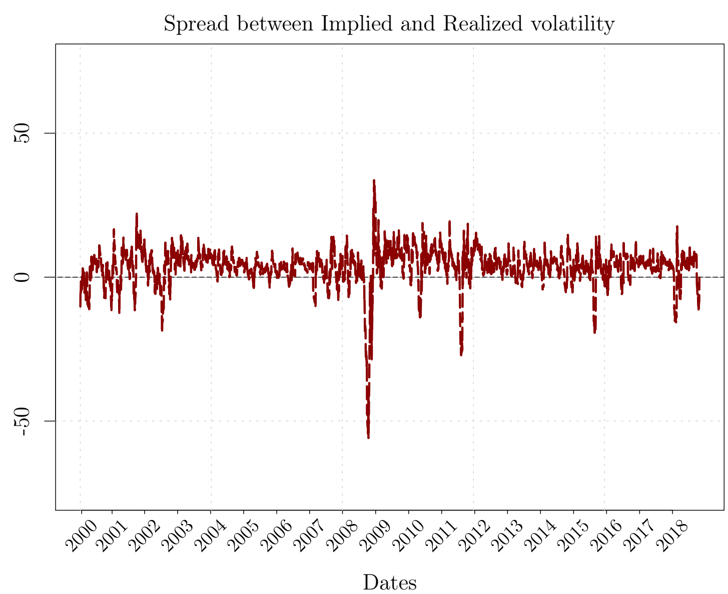

Figure 4: This plot presents the spread between the previous-month level of the VIX Index and the rolling 30-day close-to-close volatility from January 2000 to October 2018.

Electronic copy available at: https://ssrn.com/abstract=3257073

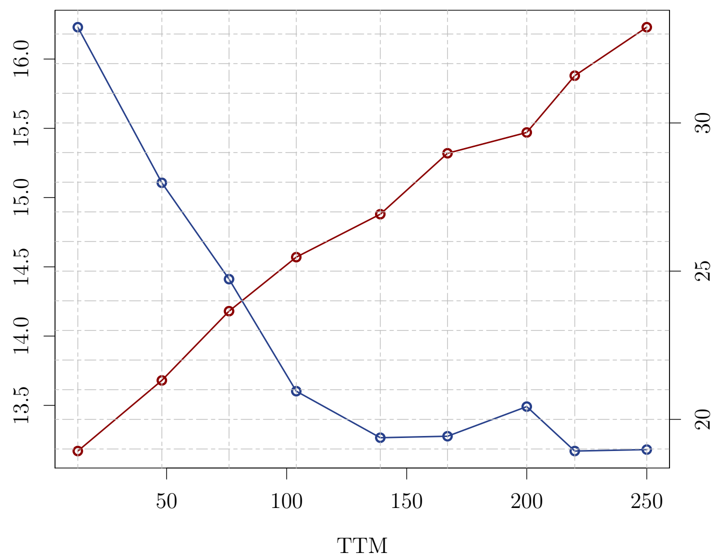

Figure 5: This plot presents the VIX futures curve before (red) and after (blue) the "volatility event" under study.

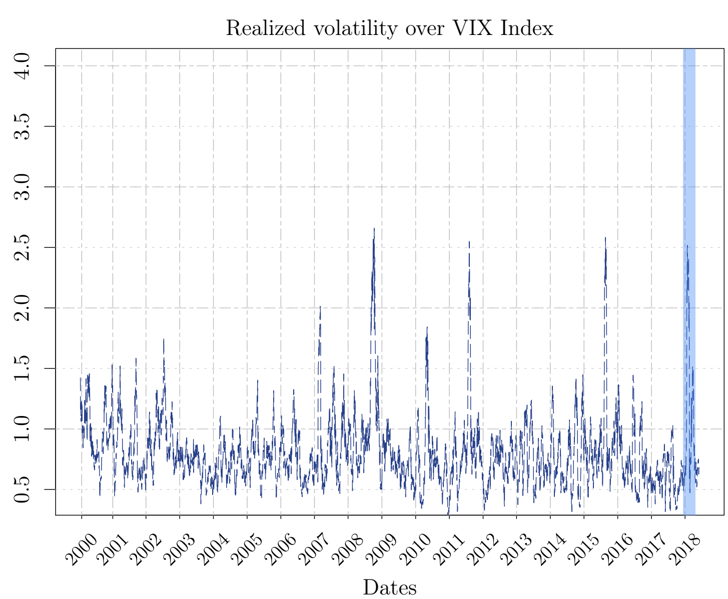

Figure 6: This plot presents the ratio between the rolling 30-day close-to-close realized volatility over the previous-month level of the VIX Index. It is possible to appreciate the "surprise factor" during the particular event at the beginning of February 2018.

Electronic copy available at: https://ssrn.com/abstract=3257073

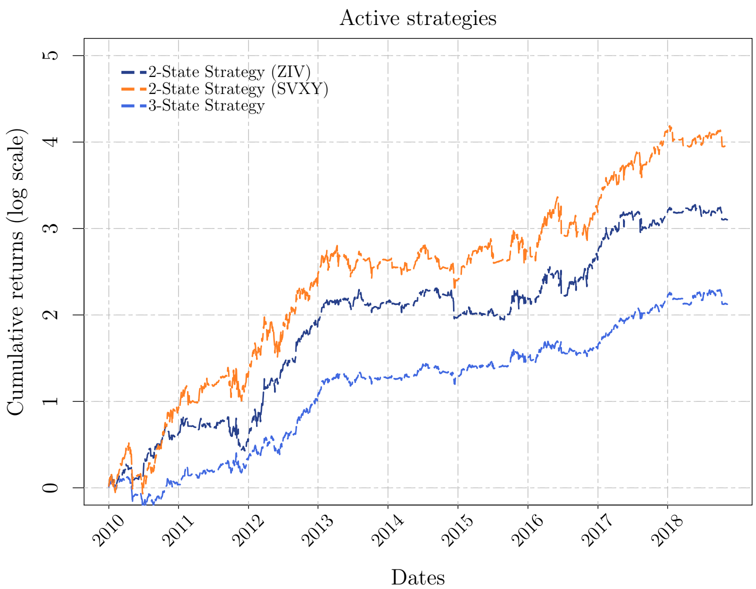

Figure 7: This plot presents the equity curve of the strategy from January 2010 to October 2018.

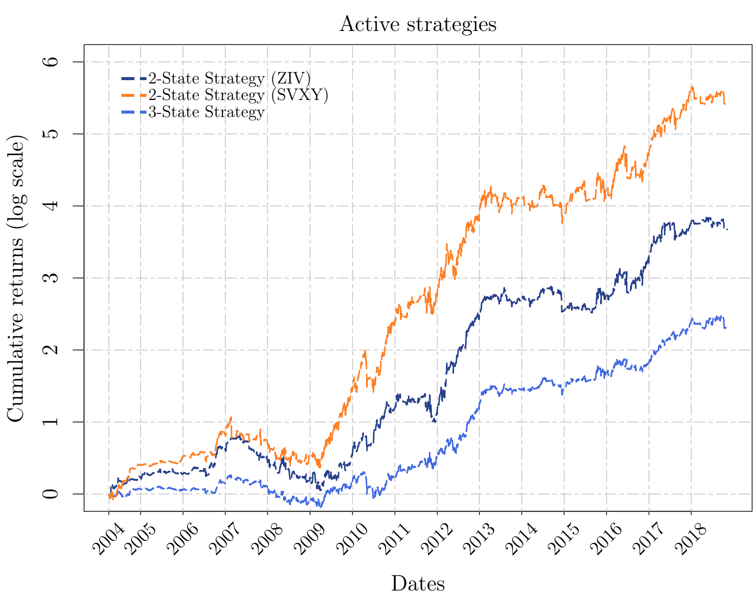

Figure 8: This plot presents the equity curve of the strategy during the whole period.

Electronic copy available at: https://ssrn.com/abstract=3257073

	S&P 500	ZIV ETN	SVXY ETF
CAGR	8.48%	10.07%	14.59%
SD	18.28%	29.48%	61.66%
Sharpe ratio	0.46	0.34	0.24%
Skewness	0.15	-0.67	-3.59%
Kurtosis	17.03	4.42	61.00%
MaxDD	-55.18%	-78.89%	-93.06%
Calmar Ratio	0.46	0.34	0.23

Table 1: This table shows the descriptive statistics and ratios on annual basis of the proposed benchmarks for the whole sample (from April 2004 to October 2018).

	S&P 500	ZIV ETN	SVXY ETF
CAGR	1.54%	-5.14%	18.32%
SD	22.73%	27.92%	49.87%
Sharpe ratio	0.06	-0.18	0.36
Skewness	0.42	-0.65	-0.94
Kurtosis	16.36	5.04	5.25
MaxDD	-55.18%	-78.89%	-92.15%
Calmar Ratio	0.03	-0.07	0.19

Table 2: This table shows the descriptive statistics and ratios on annual basis of the proposed benchmarks for the first sub-sample (from April 2004 to January 2010).

	S&P 500	ZIV ETN	SVXY ETF
CAGR	13.01%	20.79%	11.48%
SD	14.68%	30.45%	68.25%
Sharpe ratio	0.89	0.68	0.13
Skewness	-0.45	-0.69	-4.14
Kurtosis	4.39	4.09	65.91
MaxDD	-18.60%	-41.36%	-93.06%
Calmar Ratio	0.68	0.69	0.17

Table 3: This table shows the descriptive statistics and ratios on annual basis of the proposed benchmarks for the second sub-sample (from January 2010 to October 2018).

Electronic copy available at: https://ssrn.com/abstract=3257073

	2-states strategy (ZIV)	2-states strategy (SVXY)	3-states strategy
CAGR	17.24%	45.12%	28.78%
SD	21.74%	43.70%	33.20%
Sharpe ratio	0.80	1.03	0.86
Skewness	-0.60	-1.25	-0.59
Kurtosis	7.20	10.31	17.98
MaxDD	-36.37%	-50.39%	-53.45%
Calmar Ratio	0.48	0.90	0.54

Table 4: This table shows the descriptive statistics and ratios on annual basis of the proposed strategies and benchmarks for the whole sample (from April 2004 to October 2018).

	2-states strategy (ZIV)	2-states strategy (SVXY)	3-states strategy
CAGR	3.10%	28.63%	10.13%
SD	17.58%	31.82%	24.06%
Sharpe ratio	0.17	0.89	0.42
Skewness	-1.09	-2.57	-0.58
Kurtosis	12.86	22.57	9.51
MaxDD	-36.37%	-50.39%	-53.45%
Calmar Ratio	0.09	0.57	0.19

Table 5: This table shows the descriptive statistics and ratios on annual basis of the proposed strategy and benchmarks for the first sub-sample (from April 2004 to January 2010).

	2-states strategy (ZIV)	2-states strategy (SVXY)	3-states strategy
CAGR	27.00%	55.93%	41.72%
SD	24.25%	49.90%	37.95%
Sharpe ratio	1.13	1.12	1.09
Skewness	-0.48	-1.07	-0.59
Kurtosis	5.31	7.19	16.00
MaxDD	-30.63%	-44.48%	-32.60%
Calmar Ratio	0.88	1.25	1.27

Table 6: This table shows the descriptive statistics and ratios on annual basis of the proposed strategy and benchmarks for the second sub-sample (from January 2010 to October 2018).

Electronic copy available at: https://ssrn.com/abstract=3257073

	Alpha	RMRF	SMB	HML	MOM	Adj. R^2
2-state strategy (ZIV)	0.0006	0.478	0.258	0.094	0.014	0.213
	(2.93)	**(22.39)**	**(6.68)**	**(2.82)**	(0.57)	
2-state strategy (SVXY)	0.0014	1.129	0.569	0.2122	0.155	0.261
	(3.61)	**(29.91)**	**(7.87)**	**(2.76)**	**(3.06)**	
3-state strategy	0.0009	0.715	0.399	0.062	0.004	0.188
	(2.98)	**(23.82)**	**(6.95)**	(1.03)	(0.11)	

Table 7: This table shows the results of regressing the strategies returns against the four factor-model of the different strategies for the whole sample (from April 2004 to October 2018). Bolded t-stats are significant at 95% level.

Electronic copy available at: https://ssrn.com/abstract=3257073