

CAN THE VIX SIGNAL MARKET DIRECTION?

AN ASYMMETRIC DYNAMIC STRATEGY

Alessandro Cipollini

UBS Alternative Investments SGR (Italia)

Via della Spiga 30 – 20121 Milano

PhD student

University of Milan-Bicocca

Department of Mathematics and Applications

Via Cozzi 53 – 20125 Milano

Tel: +39 02 64485728

alessandro.cipollini@unimib.it

Antonio Manzini

Chief Investment Officer

UBS Alternative Investments SGR

Via della Spiga 30 – 20121 Milano

antonio.manzini@ubs.com

April 2007

Keywords and Phrases: Forecasting stock returns, Implied volatility, Asymmetric investment strategies, Market's efficiency, etc.

ABSTRACT – The article shows statistically that the VIX Implied Volatility Index is an important driver of the S&P500 future returns. The statistical analysis is performed by means of a regression based on dummy variables in order to circumvent the difficulties posed by the lack of linearity between the variables. The results obtained are then used to construct an automated procedure that signals daily whether it is convenient to invest in the S&P 500 or to stay put. Finally, we test the quality of the signal by implementing an asymmetrical buy-and-hold strategy with 3-months horizon on the S&P 500. Our results show that the strategy outperforms the long-only strategy on the same index, thus confirming a widespread belief among traders.

1. The motivation.

Implied volatility, which emerges from daily option trading activities, measures in option pricing formulas to what extent the returns of the underlying asset fluctuates from the current date until option's expiration. It is well documented that implied volatility is a reasonable forecast of future realized volatility (see for instance Poon, and Granger [2003] and Andersen, Bollerslev, Christoffersen, and Diebold [2005]) for an historical review). Surprisingly, few studies deal with a similar issue, namely the possible relationship between implied volatility and future stock returns. Yet, market's participants, and in particular traders, are well aware of it. Indeed, a widespread belief among them holds that swings in implied volatility yield good clues on future directions of the market. An increase in the implied volatility value is associated with fear in the market, whereas a decline indicates complacency. As a measure of fear and complacency, implied volatility is often used as a contrarian indicator: prolonged and/or extremely high VIX readings indicate a high degree of anxiety – or even panic – among traders, and are regarded as a bullish indicator. Prolonged and/or extremely low readings indicate a high degree of complacency, and are generally regarded as a bearish indicator.

Can this belief find a reasonable explanation? There are two possible answers. Firstly, a *statistical answer*, which simply tries to assess with the appropriate statistical tools whether there is a significant statistical link between current levels of implied volatility and the future stock returns. Secondly a *theoretical answer*, which endeavours to identify along which channels the implied volatility levels are connected with the future stock returns.

We provide, in this paper, an answer to the statistical issue, by adopting the procedure of Giot [2005] and of Campbell and Shiller [1998]. For what concerns the

theoretical answer, we content ourselves with the insights provided by Capital Arbitrage Pricing Model (CAPM), which seems to offers an immediate answer:

$$R_p - r_f = \beta(R_M - r_f) = \rho \frac{\sigma_P}{\sigma_M} (R_M - r_f) = \sigma_P \rho S R_M.$$

where $(R_p - r_f)$ is the risk premium of asset (or portfolio) P ; β is P 's systematic risk relative to market M ; ρ is the correlation between returns from P and M , $(R_M - r_f)$ is the market's risk premium and σ is volatility (annualized standard deviation of returns).

The higher volatility translates into a high risks premium (excess return). The issue is that σ_P is an instantaneous unobservable volatility. The idea of adopting the implied volatility instead of the historical realized volatility solves the problem, but leaves the door open for further theoretical investigation justifying their substitution in the CAPM formulation. Indeed, while the implied volatility is directly linked to the options' prices, the link to the underlying stock prices is indirect and in particular non-linear, thus difficult to investigate. However the rationale in practical terms is easily inferred: increasing implied volatility is brought about by the rise of options' prices, which could be associated with the desire of hedging the equity risk.

In our study, we adopted as a measure of the implied volatility the CBOE (Chicago Board Options Exchange) volatility index (VIX), and as future stock's returns the 3-months S&P 500 Index returns¹. The first part of the paper is devoted to the investigation of a statistical significant relationship between the VIX Implied Volatility Index prices and the 3-months S&P 500 Index returns. Once established the statistical significance of such a relationship, the attention has been focused on developing a quantitative procedure able to signal daily, according to the implied volatility level, the

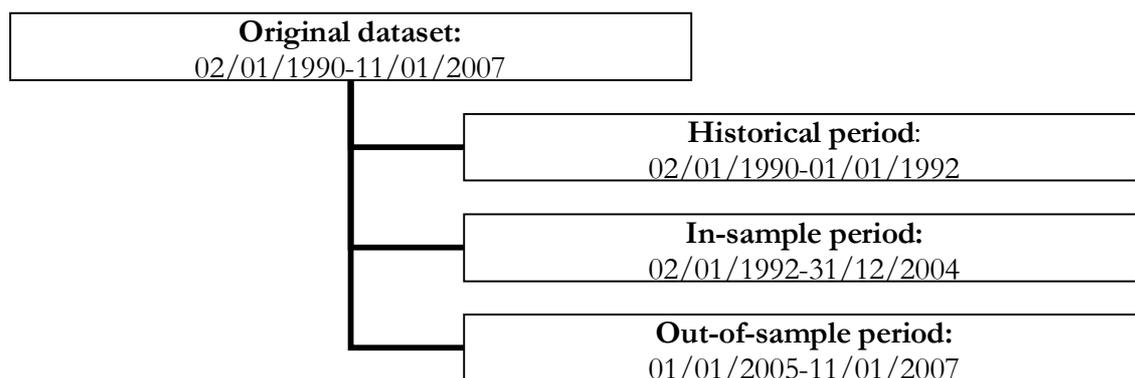
¹ It is clear that the period of forward S&P 500 returns is arbitrary. As asset managers we are interested in a long term investment return. We feel that it is reasonable to limit our choice to a 3 months period.

expected market's direction on a 3-months horizon and consequently the convenient times for opening a buy-and-hold strategy (long position). Finally in order to have a sound grasp of the signal efficiency, a simple buy-and-hold trading strategy, driven only by the implied volatility signal, has been back tested. In particular the strategy has been analysed in comparison to a long always investment.

2. Analysis of the dataset used.

The dataset used here consists of a period of 16 years that goes between the 02/01/1990 and the 11/01/2007 on a daily basis (trading days) and therefore in 4433 observations. The variables observed are the CBOESXP (NEW) VIX-PRICE INDEX and the S&P 500 COMPOSITE-PRICE INDEX series in the original currency (USD)².

In order to preserve an out-of-sample period sufficiently long for back testing the relationship, the dataset has been divided into these three periods:



The dataset considered contains sufficiently different market environments in order to be, in our opinion, a representative sample of analysis (see the graph in exhibit I). A

² The provider used is Datastream.

comparison between the VIX Implied Volatility Index with its long term moving average (24-months moving average³), contributes to signal the high/low volatility environment.

It is useful to compute and represent (see the graph in exhibit II) the series of 3-months forwards S&P 500 Log>Returns. At a given time t of the financial horizon, the S&P500 3-months forward log>Returns are defined as:

$$3^{mths} r_{S\&P500,t} = \ln(S \& P500_{t+64} / S \& P500_t) \quad \forall t$$

3. The VIX Implied Volatility Index does signal the future market's behaviour?

While it is clear that negative returns are associated with increased implied volatility, there is a growing debate as how implied volatility indexes can indicate overbought or oversold market's conditions. The main difficulty lies in the fact that the relationship between implied volatility and the option's underlying spot price fails to be linear. In particular a linear regression analysis of the relationship between VIX Implied Volatility Index and the 3-months forward future log>Returns does not seem to offer interesting results. In statistics there are many techniques through which it is possible to handle non-linearly related regressors with a dependent variable. In our case, following the original approach due to Giot [2005] and Campbell and Shiller [1998], we adopted a dummy variables technique.

Here the procedure followed is based on four principal steps:

1. **Determining percentiles.** We subdivide each of the 24-month VIX rolling windows of our sample into 22 percentiles, i.e. $P_{1t}, P_{2t}, \dots, P_{21t}, P_{22t}$.

³ The moving average has been computed on a rolling period of 24-months, on the period of time between the 02/01/1990 and the 11/01/2007. Alternative choices can be made in order to measure long term volatility (such as with Garch models) but it is outside the scope of the article.

2. **Ranking position.** We determine in which of the 22 classes defined by the percentiles the current VIX level falls. We then translate this position into a rank (R_t), such that at the first class corresponds rank one, at the second rank two and so on:

$$R_t = i \Leftrightarrow P_{it} \leq VIX_t < P_{(i+1)t}.$$

3. **Coding variables.** We code the dummy variables by means of the current, current VIX rank:

$$\begin{cases} D_{it} = 1 \text{ if } R_t = i \\ D_{it} = 0 \text{ if } R_t \neq i \end{cases}, 1 \leq i \leq 22.$$

4. **Regression.** Estimation of the 22 dummy variables' coefficients, yielding the expected returns:

$$\begin{cases} \hat{\alpha}_i : 1 \leq i \leq 22 \\ 3^{mths} r_{S\&P500t} = \sum_{i=1}^{22} \hat{\alpha}_i D_{it} + \varepsilon_t \end{cases}$$

Specifically, we rank observed implied volatility into one of the 22 equally spaced percentiles of the VIX Implied Volatility Index observed at any time on a rolling period of 24-months⁴. In particular note that at time t the 22-percentiles are computed by considering the VIX Implied Volatility Index values, starting from 24-months (510 trading days) in the past up to the present time: $VIX_{t-509}, \dots, VIX_{t-1}, VIX_t$. Then VIX_t is compared to these 22 equally spaced percentiles and ranked accordingly, say, in position R_t . In the cases in which VIX_t is lower than the minimum of all the 510 past values it will be ranked in class 0 and; similarly; if higher than the maximum of all past values in class 22. Next, according to the position of implied volatility, a dummy variable D_{it} is

⁴ By equally spaced 22-percentiles is meant the 4.55%, 9.09%, 13.64%,...95.45%,100%-percentiles. The number of classes (22) has been chosen empirically in order to have a "good" fit of all the ranking classes with the data, in particular to avoid concentration in one specific class. The choice of a 24 months window is arbitrary, yielding nevertheless a stable scale for the ranking of volatility.

defined such that $D_{it} = 1$ if $R_t = i$ and $D_{it} = 0$ otherwise. Thus the R_t -classification is mapped into 22 distinct dummy variables that can be used in a linear regression model. This procedure is repeated for the whole in-sample period (01/01/1992-31/12/2004). The graph in exhibit 3 shows the behaviour of the implied volatility in the period of interest, in terms of ranking classes.

The linear regression model estimated is:

$$3^{mths} r_{S\&P500,t} = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \dots + \alpha_{21} D_{21,t} + \alpha_{22} D_{22,t} + \varepsilon_t, \quad \forall t$$

where $D_{1,t}, D_{2,t}, \dots, D_{21,t}, D_{22,t}$ are the dummy variables and ε_t is the residual term. The coefficient can be directly interpreted as the expected return for the given time horizon when VIX_t is ranked in category R_t at time t . For example the estimated $\alpha_{20,t}$ for the S&P500 Index at the 3-months time horizon gives the expected return for the S&P500 Index whenever the VIX Implied Volatility Index is ranked in 20. From the statistical point of view the estimates have been obtained through the OLS procedure. Moreover, in order to improve the efficiency of the model, the estimates have been corrected for the heteroskedasticity effect by the Newey-West procedure. The table in exhibit 4 reports the estimates obtained together with their statistical significance

The estimates obtained are all different from zero with an error probability smaller than 5%, except for the dummy variables, which correspond to a low implied volatility rank. This fact suggests that there exists a meaningful relationship between VIX Implied Volatility Index and the 3-months forward S&P500 future Log>Returns. In particular the expected returns are high for higher classes of implied volatility and become low and sometimes negative for lower classes of implied volatility⁵. These results lend support to the hypothesis that high levels of implied volatility signal attractive entry

⁵ It is reasonable to expect that coefficients that correspond to lowest classes of implied volatility are negative. Moreover it has to be stressed that in the case here analysed the estimates provided for those classes of volatility are not statistically meaningful.

points for a long position. In particular the higher the spike of implied volatility the higher the return on a 3-months buy-and-hold position on the S&P500.

4. A simple Buy-and-Hold trading strategy driven by the VIX signal.

On these results, we build an elementary trading strategy that simply opens a buy-and-hold position on the S&P 500 Index whenever the estimated expected returns (the regression coefficients) have to be considered interesting for investing purposes⁶.

As the future return on the S&P 500 cannot be explained only by its current implied volatility, we acknowledge a lot of noise in the regression residuals⁷. To tackle the problem, we run the original regression by modelling the residuals with an autoregressive component of order 1. The resulting new regression is not linear:

$$\begin{cases} 3^{mths} r_{S\&P500,t} = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \dots + \alpha_{21} D_{21,t} + \alpha_{22} D_{22,t} + \varepsilon_t, & \forall t, \\ \varepsilon_t = \phi_1 \varepsilon_{t-1} + u_t \end{cases}$$

for u_t residual term. The evaluation's problem of the coefficients' estimates may be tackled in different manners: we have chosen to substitute the second equation in the first one and to perform the Marquardt non-linear least square algorithm. The estimates are reported in the table in exhibit 5. We consider as interesting the volatility signals ranked from the 3th class upward (dummy 3 and above). As the graph in exhibit VI shows, the rationale behind the choice of the coefficients hinges on the fact that we require them to be positive, since we are interested in expected positive returns.

Our strategy consists in opening daily a 3 months buy-and-hold position on the S&P 500 whenever the ranking position of the implied volatility level falls into class 3

⁶ Despite its practical relevance, this is a paper strategy. Hence, transaction costs are not considered.

⁷ The residual graph together with other statistical details can be provided upon request.

and above, those with positive expected returns. In the remaining cases, i.e. when the signal falls into classes 0, 1 and 2, the strategy invests in a risk-free asset⁸. The graph in exhibit 7 and the accompanying table show a comparison between the series of log-returns of a long always 3-months position and the series of log-returns of the buy-and-hold strategy driven by the signal. It is interesting to see that the back-test validates the strategy since most of the time it captures the upside of the long-only position and provides an hedge against periods with negative returns. Indeed, the pay-off is intended to be asymmetrical: whenever the signal is effective the downside is protected at the Libor rate and the upside is unbounded and coincides with the upside of a long-always 3-months position (as stylized in the graph of exhibit 8, grey line). When the reliability of the signal breaks down, 2 possibilities arise: either the VIX signal suggests entering when the S&P 500 goes down or the signal suggests staying still when the market goes up (graph of exhibit 8, broken line).

In reality, our strategy pay-off turns out to be asymmetrical and when it fails to be so, in case of signal break down, resulting losses are limited and do not offset the appealing gain (see the graph and the table in exhibit 9).

Is the strategy beating the S&P 500⁹? The answer is affirmative, as results are superior as shown in the graph of exhibit 10, which reports the cumulative returns of the long-always and the buy-and-hold when coefficients are positive. We suspect that even better results could be achieved if we could leverage the investment according to the size of the forecasted expected returns represented by the beta coefficients (see graph of exhibit 6).

⁸ As a risk-free asset here is taken a risk-free bond with as spot interest rate the standard Libor rate on a daily basis.

⁹ The cumulative returns of both the S&P 500 and the long always strategy, the benchmark or our strategy, are equivalent.

5. Conclusions

Can the VIX signal market direction? Yes, it can. However, our research indicates that *the signal is loud and clear* when the implied volatility is high (or, even better, when it spikes). On the contrary, at low levels of implied volatility the model is less effective.

Moreover the out-of-sample back testing demonstrates *the ability of the model to create value by timing the optimal entry points for a long position.* The strategy demonstrates a *higher return, lower volatility and, consequently, a higher Sharpe ratio* compared to a continuous long-only investment in the S&P 500.

The model also represents an *independent source of alpha*, which is an important part of alpha diversification. Portfolios driven by other factors like value, growth, or fundamentals are thus likely to increase their Sharpe ratio when allocating part of their equity investments to the VIX buy-and-hold strategy.

6. References

Poon S.H. and Granger C.W.L., "Forecasting Volatility in Financial Markets: A Review.", *Journal of Economic Literature*, Vol. XLI , pp.478-539, 2003.

Andersen T. G., Bollerslev T., Christoffersen P.F. and Diebold F.X., "Volatility Forecasting", *Working Paper, National Bureau Of Economic Research*, Cambridge, 2005.

Heydon T., "Fear and Greed in Global Asset Allocation", *Journal of Investing*, Vol. 9 Issue 1, pp. 27-35, 2000.

Giot P., "Relationships between Implied Volatility Indexes and Stock Index Returns. Are implied volatility indexes leading indicators?", *Journal of Portfolio Management*, pp. 92-100, 2005.

Campbell, J. Y., and Shiller J. R.. "Valuation Ratios and the Long-Run Stock Market Outlook." *The Journal of Portfolio Management*, 24, pp. 11-26, 1998.

Quandt R.E., "Computational Problems and Methods", Chapter 12 in Z. Griliches and M.D. Intriligator (eds.), *Hanbook of Econometrics*, Vol 1, North-Holland, 1983

Exhibit I

The implied volatility and the market's environment.

In the graph is reported the behaviour of the VIX CBOE's Implied Volatility Index (NEW) together with the S&P 500 Composite Price Index, regarding to the whole dataset period (02/01/1990-11/01/2007).

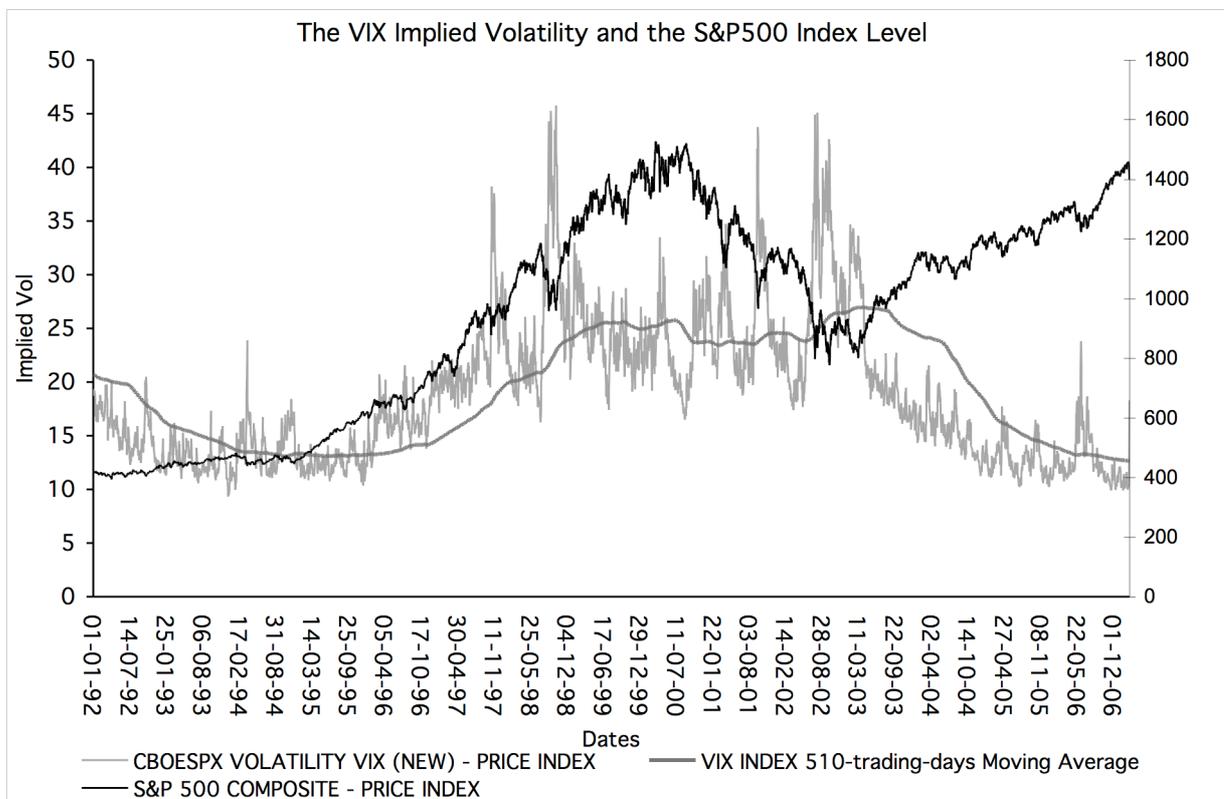


Exhibit II

The 3-months forward Log Returns.

In the graph the log-returns of a simple buy-and-hold position on the S&P 500 Index are compared with the VIX Implied Volatility Index. Note that a high period volatility environment seems to boost the oscillations of the log-returns. The data concerns the whole dataset period (02/01/1990-11/01/2007).

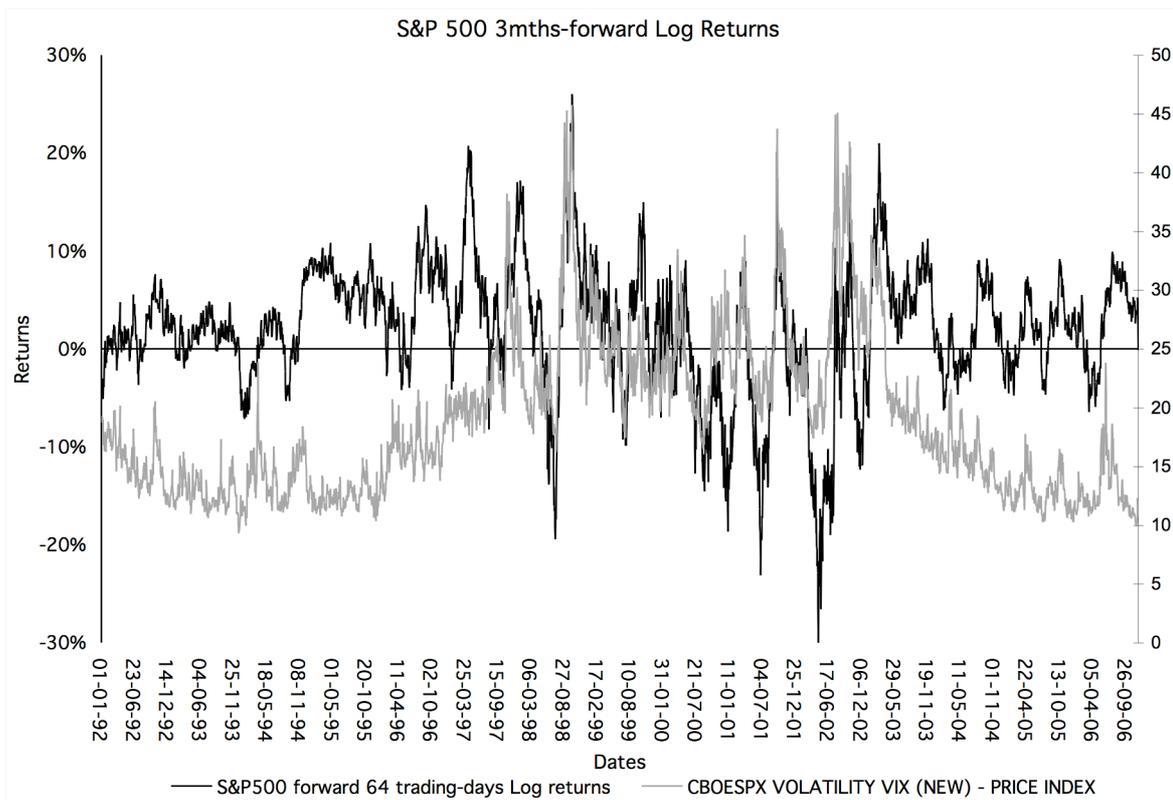


Exhibit III

The VIX Implied Volatility Index ranked in class.

In the graph are reported the VIX Index Implied Volatility ranked in the 22 classes, referring to the in-sample period (02/01/1992-31/12/2004). The implied volatility spans all the classes considered. In particular the frequency of cases when implied volatility is in the 22nd class is 0.82% and in the 0th class 12.4%.

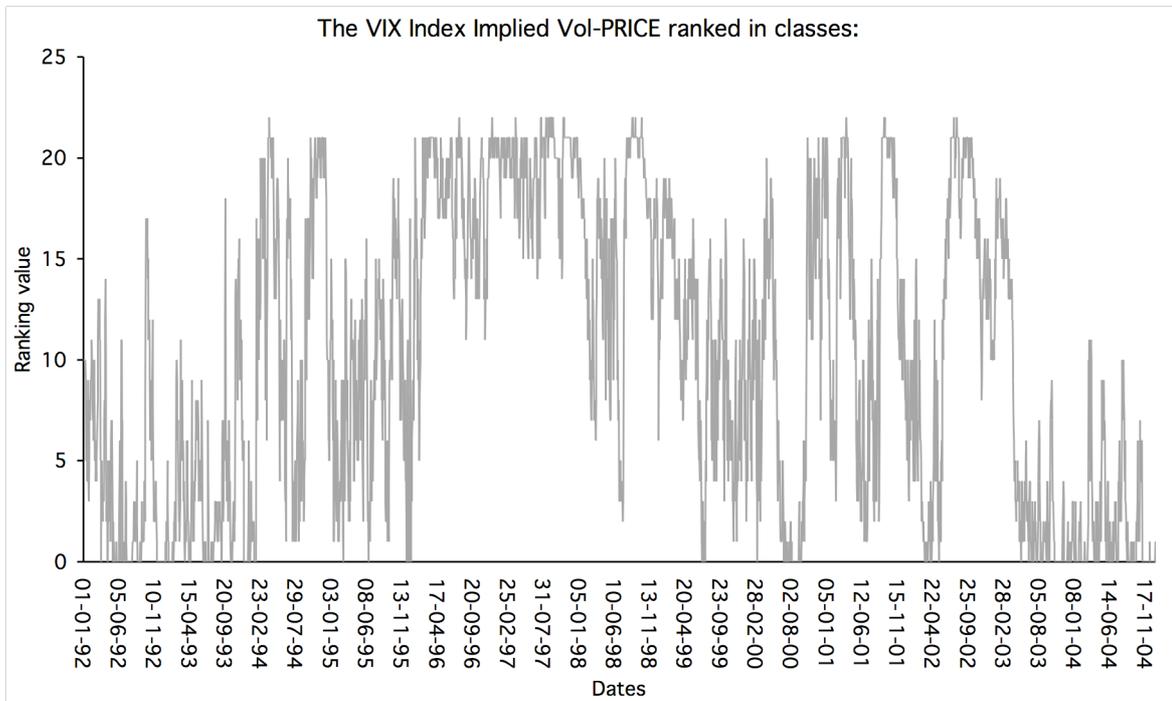


Exhibit IV

The estimates obtained.

The table reports the coefficients estimated of the model on the in-sample period (02/01/1992-31/12/2004):

$$3^{mths} r_{S\&P500,t} = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \dots + \alpha_{21} D_{21,t} + \alpha_{22} D_{22,t} + \varepsilon_t, \quad \forall t,$$

together with some elementary statistics. Note that the estimates in bold are statistically different from zero with an error of 5%.

The coefficients have to be interpreted as 3-months log-returns of a buy-and-hold position on the S&P500 Index.

Dependent Variable: SXP500K				
Method: <i>Least Squares</i>				
Sample: <i>1/02/1992 12/31/2004</i>				
Included observations: <i>3393</i>				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DUMMY0	0,35%	0,32%	1,11	26,64%
DUMMY1	-0,27%	0,46%	-0,60	55,12%
DUMMY2	-0,44%	0,51%	-0,86	38,74%
DUMMY3	0,09%	0,51%	0,17	86,79%
DUMMY4	-0,16%	0,56%	-0,28	77,86%
DUMMY5	-0,57%	0,61%	-0,93	35,26%
DUMMY6	1,16%	0,60%	1,94	5,25%
DUMMY7	1,32%	0,64%	2,05	4,01%
DUMMY8	1,28%	0,65%	1,97	4,90%

DUMMY9	1,85%	0,62%	3,00	0,27%
DUMMY10	0,93%	0,61%	1,52	12,92%
DUMMY11	2,11%	0,64%	3,27	0,11%
DUMMY12	0,81%	0,65%	1,24	21,44%
DUMMY13	2,49%	0,60%	4,11	0,00%
DUMMY14	1,52%	0,59%	2,58	0,99%
DUMMY15	2,57%	0,60%	4,28	0,00%
DUMMY16	3,67%	0,59%	6,28	0,00%
DUMMY17	2,72%	0,55%	4,93	0,00%
DUMMY18	4,01%	0,55%	7,32	0,00%
DUMMY19	3,50%	0,54%	6,46	0,00%
DUMMY20	4,33%	0,48%	9,01	0,00%
DUMMY21	6,38%	0,37%	17,11	0,00%
DUMMY22	8,71%	1,24%	7,05	0,00%
R-squared	0,095	Mean dependent var		0,019
Adjusted R-squared	0,089	S.D. dependent var		0,068
S.E. of regression	0,065	Akaike info criterion		-2,609
Sum squared resid	14,422	Schwarz criterion		-2,567
Log likelihood	4449,54	Durbin-Watson stat		0,053

Exhibit V

The estimates obtained.

The table reports the coefficients estimated of the model on the in-sample period (02/01/1992-31/12/2004):

$$\begin{cases} 3^{mths} r_{S\&P500,t} = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \dots + \alpha_{21} D_{21,t} + \alpha_{22} D_{22,t} + \varepsilon_t, \quad \forall t, \\ \varepsilon_t = \phi_1 \varepsilon_{t-1} + u_t \end{cases}$$

together with some elementary statistics. We do not report t-statistics since being the regression non-linear they do not have any statistical meaning. Note the high R^2 .

Dependent Variable: <i>SXP500K</i>				
Sample: <i>1/02/1992 12/31/2004</i>				
Included observations: <i>3393 after adjustments</i>				
Convergente: <i>achieved after 5 iterations</i>				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DUMMY0	-0,58%	1,11%	-0,52	12,47%
DUMMY1	-0,14%	1,11%	-0,13	13,56%
DUMMY2	-0,03%	1,11%	-0,02	22,35%
DUMMY3	0,24%	1,11%	0,22	3,54%
DUMMY4	0,39%	1,11%	0,36	4,67%
DUMMY5	0,56%	1,11%	0,50	6,10%
DUMMY6	0,83%	1,11%	0,75	4,54%
DUMMY7	1,02%	1,11%	0,92	3,67%
DUMMY8	1,34%	1,11%	1,20	2,28%
DUMMY9	1,46%	1,11%	1,32	1,82%

DUMMY10	1,65%	1,11%	1,49	1,32%
DUMMY11	1,70%	1,11%	1,54	1,25%
DUMMY12	1,97%	1,11%	1,77	7,63%
DUMMY13	2,24%	1,11%	2,02	4,32%
DUMMY14	2,53%	1,11%	2,29	2,24%
DUMMY15	2,90%	1,11%	2,61	0,90%
DUMMY16	3,05%	1,11%	2,75	0,60%
DUMMY17	3,41%	1,11%	3,07	0,21%
DUMMY18	3,85%	1,11%	3,47	0,05%
DUMMY19	4,27%	1,11%	3,84	0,01%
DUMMY20	4,80%	1,11%	4,31	0,00%
DUMMY21	6,00%	1,12%	5,38	0,00%
DUMMY22	7,31%	1,14%	6,44	0,00%
AR(1)	97,9%	0,35%	281,67	0,00%
R-squared	0,962	Mean dependent var		0,019
Adjusted R-squared	0,962	S.D. dependent var		0,068
S.E. of regression	0,013	Akaike info criterion		-5,801
Sum squared resid	0,592	Schwarz criterion		-5,757
Log likelihood	9865,895	Durbin-Watson stat		2,040
Inverted AR Roots	0,98			

Exhibit VI

The regression coefficients.

The table reports the coefficients estimated of the model on the in-sample period (02/01/1992-31/12/2004):

$$\begin{cases} 3^{mths} r_{S\&P\ 500,t} = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \dots + \alpha_{21} D_{21,t} + \alpha_{22} D_{22,t} + \varepsilon_t, \quad \forall t, \\ \varepsilon_t = \phi_1 \varepsilon_{t-1} + u_t \end{cases}$$

that has to be interpreted as quarterly 3-months log returns of a buy-and-hold position on the S&P 500 Index.

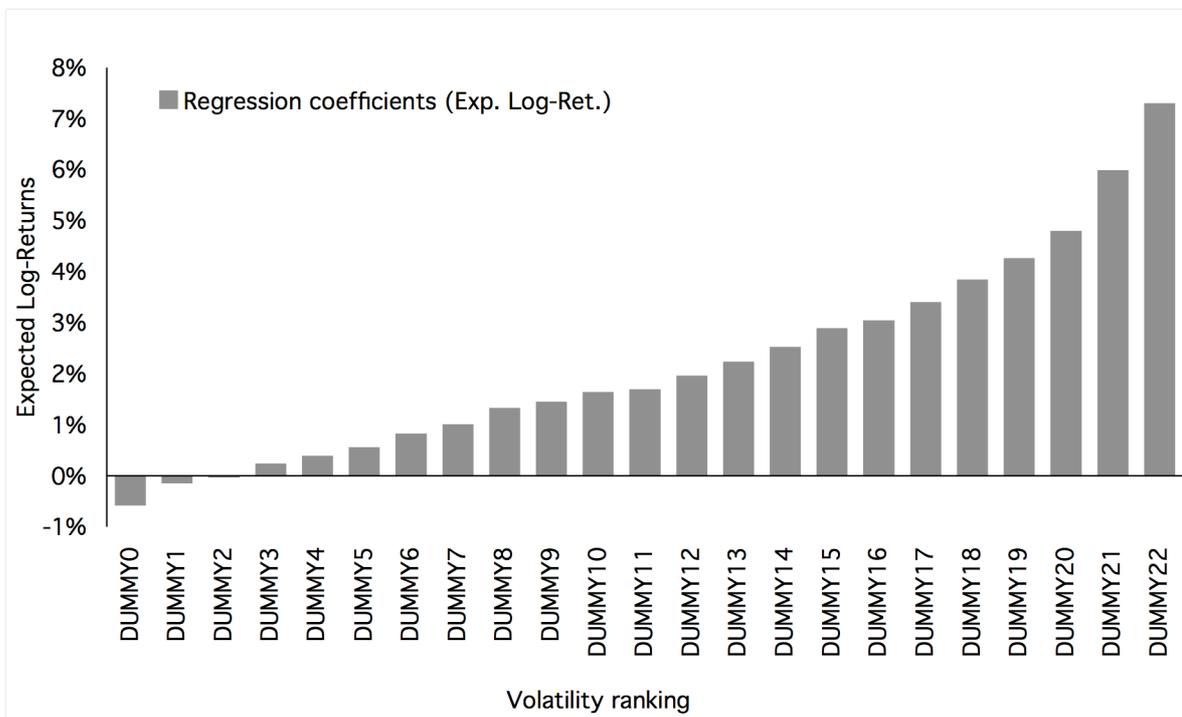
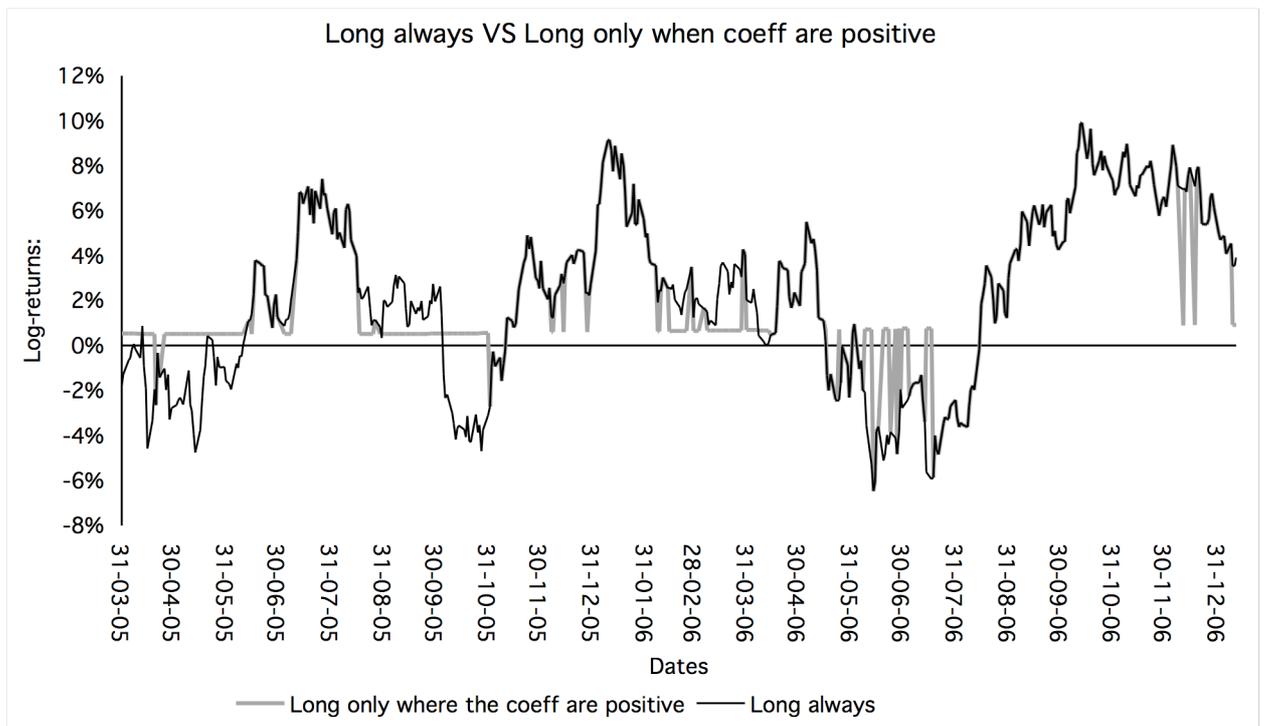


Exhibit VII

The buy-and-hold strategy when coefficients are positive.

The graph is a comparison among the long-always position with the buy-and-hold strategy driven by the VIX signal. Note that all the upside of the long-always is captured while the downside is almost capped at the Libor rate. In the subsequent table are reported the annualized returns, the volatilities and the corresponding Sharpe ratios.



	Long always	Buy-and-hold when coeff >0
Annualized return	9,80%	10,18%
Annualized volatility	13,89%	11,90%
Sharpe ratio	0,458	0,626

Exhibit VIII

The buy-and-hold strategy when coefficients are positive.

The graph is the ideal pay-off of the buy-and-hold strategy driven by the VIX signal implemented. Note that the latter pay-off resembles a call option having as underlying the long-always position on the index.

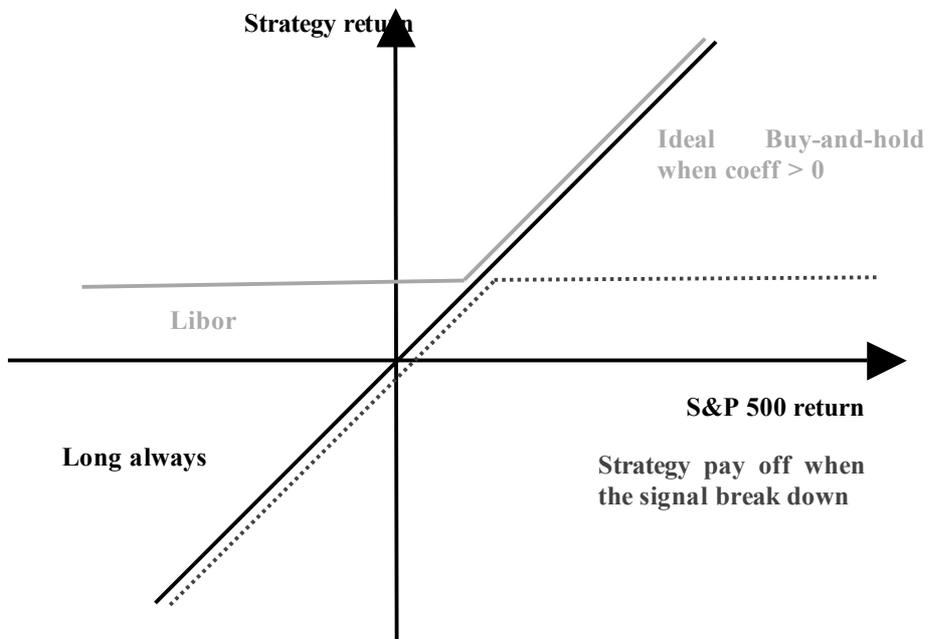
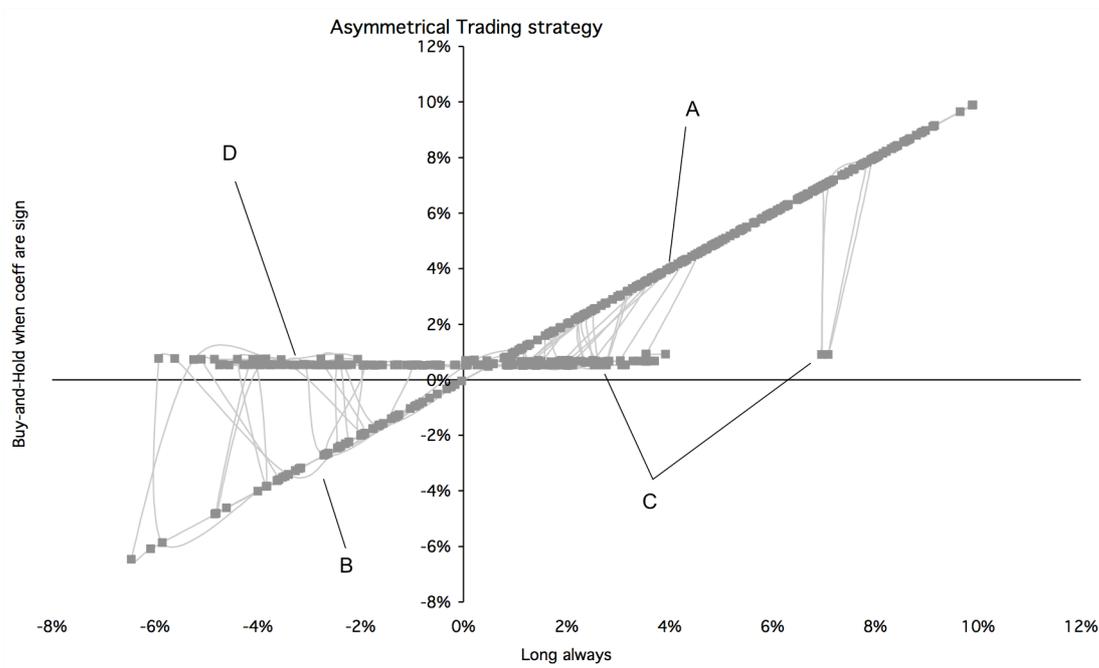


Exhibit IX

The real strategy profile.

The graph is the real pay-off of the buy-and-hold strategy. In the subsequent table are reported the annualized returns and the frequencies (in brackets) of the four possible situations. In particular points on segments A and D correspond to the intended asymmetric return profile of the strategy, while points on the segments B and C to the break down of the signal



		Long always	
		positive returns	negative returns
Buy-and-hold when coeff are >0	On	22,01% A (235)	-9,18% B (59)
	Off	-0,52% C (95)	13,5% D (77)

Exhibit X

The cumulative returns.

The graph shows a comparison between the cumulative returns of the Long-always and the buy-and-hold cumulative returns when coefficients are positive. Moreover notice that in the graph is represented also the VIX level rescaled of three months (at the current returns correspond the implied volatility of three months earlier). Significant improvements in the cumulative returns of the VIX driven strategy, are seen in correspondence of peaks of the implied volatility.

