Trading Bull- and Bear-Markets with a Hidden Markov Model.

Chrilly Donninger
Chief Scientist, Sibyl-Project
Sibyl-Working-Paper, April 2017

"There is only one side of the market and it is not the bull side or the bear side, but the right side".
(Jesse Livermore).

Abstract:
I developed a Hidden-Markov-Model for tail-risk-protection of the SPY with VIX-Futures [1] in the previous working paper. This working paper is a follow-up. It applies the model to different equity-market indexes and for switching between equity- and treasuries ETFs. It considers additionally the effect of leveraged ETFs. The analysis extends the rather promising results of the original paper.

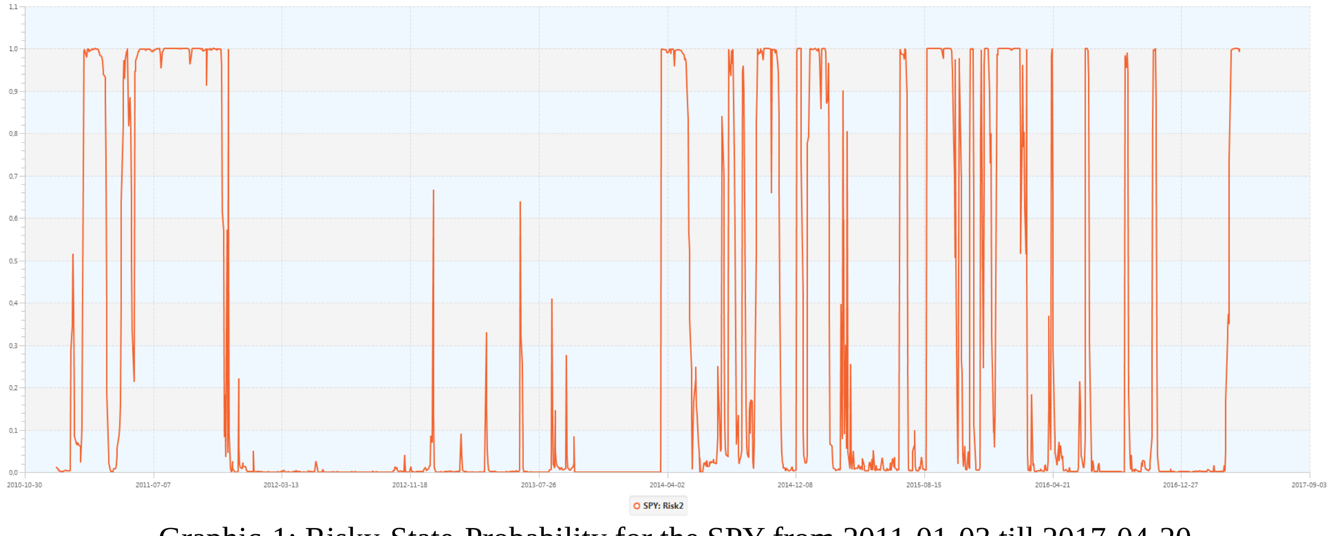
Introduction:
There are papers which show the profitability of protecting an equity market portfolio with VIX-Futures. But they only consider the period of the 2008 crash. It is a trivial result that an insurance is beneficial in case of a damage. There is a general consensus that paying the insurance constantly premium is too costly. One has to find a reasonable working classification of bull- and bear-markets. Buying protection in a bull-market is a waste of money.
A well known statistical tool to address such a question is the Hidden-Markov-Model (HMM). The method was initially developed in the field of temporal pattern recognition such as speech, handwriting or gesture recognition.
The assumptions of the model are: The distribution is i.i.d for a given – non observable – state, but differs significantly between states. The observed variable is usually correlated. The correlation is a result of state transitions. The state-transitions depend only on the last state (this is the Markov-condition) and not on the full state-space history. The Markov-condition simplifies greatly the estimation of the hidden latent states. The number of states must be specified beforehand. The observed sequence in the original applications are discrete variables. But it is relative straightforward to extend the model to a continuous distribution. It is usual to assume a Multivariate Normal-Distribution. This assumption can be justified by the fact that almost any (reasonable) distribution can be approximated by a mixture of Gaussian's. But this is only a theoretical result. The number of states would usually be too large to be of any practical use. According the motto "All models are wrong, some are useful" one is interested in calculating in a reliable way the latent states and does not care too much about the Gaussian assumption and the Markov-condition.

The model developed in [1] uses a Multivariate-Normal with 2 dimensions. One can select different variables to infer the hidden states. The by far best results were by using the daily-absolute return (scaled by 100*sqrt(252.0)) of the protected equity-market ETF (e.g. the SPY) and the VIX-Future Term Structure. The VIX-Futures Term-Structure is the difference in price between the 2nd and 1st Future. The 1st Future is rolled over 5 trading days before expiry. This avoids the erratic price behavior of the 1st Future immediately before expiry. The Term-Structure is usually positive, the Futures are in contango. The price difference is typically 1.0. The Futures are only during a market-turmoil in backwardation.

Implementation of the HMM:
The parameters of the HMM are calculated with the Baum-Welch Algorithm. The algorithm is conceptually simple, but handling the numeric problems of over- and underflow is not trivial.

I ported the C++ implementation of Press [2] to Java. The implementation of Press assumes a discrete distribution. It is straightforward to replace this term by the density of the 2D-Normal. But the implementation failed when I tried also a plain Normal-Distribution. The algorithm selects for one state a single observation and classifies the other measurements to the remaining state. This single observation has a variance of zero and hence an infinite density. This is a known problem which can be addressed by regularization terms [3]. I did not encounter this problem for the 2-dimensional case.
For trading purposes one recalculates for each day the HMM. One uses always a sliding window of 2 years (504 trading-days). I tried a HMM with 2 and 3 states. The 2 states model is clearly superior. There are – for the model – only bears and bulls. But the market can be in a transitory phase in between. The Baum-Welch algorithm calculates for the whole time-range the daily state-probabilities. For trading only the last entry is of interest. The algorithm has of course no notion of bulls and bears. It classifies the trading states to maximize the overall likelihood. The numbering of the states is arbitrary. It happens that the risky bear-state is at time T state-0 and at time T+1 state-1. A state k (with k=0,1) is defined risky, if the state-marginal mean of the absolute return of k is greater than of the marginal mean of the other state and the marginal-mean of the VIX-Futures Term-Structure is less than in the alternative state. In other words: The risky-state has a larger volatility and the Futures are in backwardation.
There are a few occasions where e.g. state 1 has the greater mean absolute return but also the greater mean VIX-Futures Term-Structure. This can happen if there was no major market-turmoil within the last 2 years. There was in fact only one market regime. But the model always tries to find two and places the measurements rather arbitrarily into the state-bins. This situation was resolved by setting the risky-state-probability to Zero. Theoretically it could also be a crash lasting for 2 years. But this is an extremely unlikely situation. Using a 2 years window means that the definition of risky, the distribution of the absolute-return and of the Term-Structures differs from time period to time-period. It is a relative measure: Risky in terms of the last 2 years.

Graphic-1 shows for the SPY the probability of being in the risky-state. The transition is usually very fast from a probability close to Zero (a bull-market) to close to One (bear-market). This holds also in the other direction. One sees on the left the relative long lasting market-turmoil in the summer of 2011. The other risky-regimes are of shorter duration.



Graphic-1: Risky-State-Probability for the SPY from 2011-01-03 till 2017-04-20

Dimension-2, the VIX-Futures Term-Structure, depends on the behavior of the S&P-500. The model uses this variables also for the QQQ (Nasdaq-100), DIA (Dow-Jones) and the IWM (Russel-2000).

There are own volatility-indexes available. But no Futures are traded on these alternative indexes. The 1st dimension measures the realized volatility of these indexes. The indexes are highly correlated. Hence the Risk-Probabilities are quite similar. Only the IWM differs in between – from April 2014 to April 2016 - significantly from the SPY. The small caps performed in this period much worse.

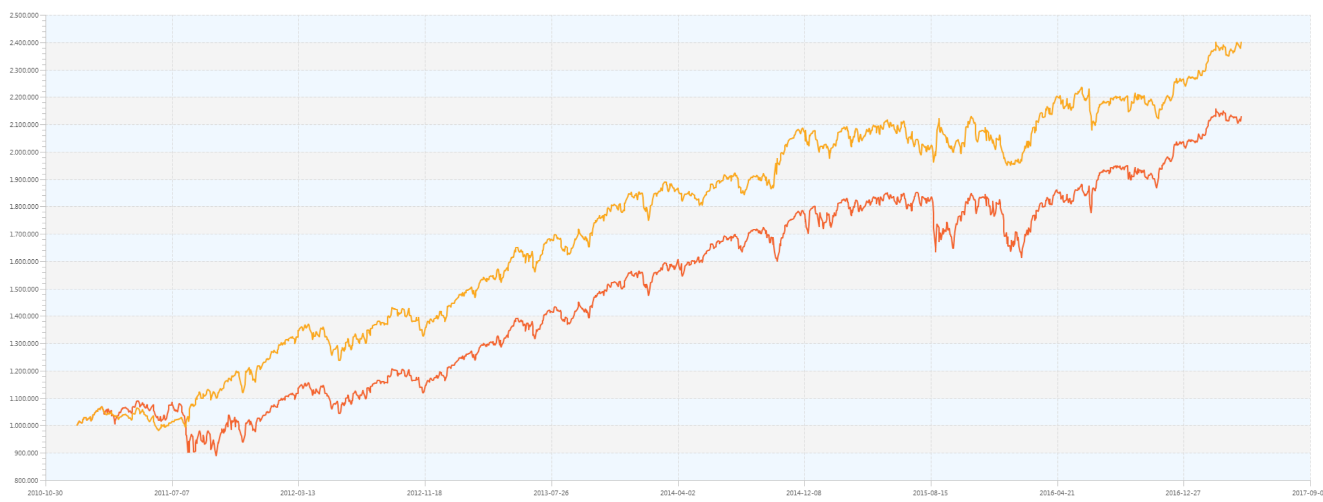
The Tail-Protection-Strategy:

One buys initially at 2011-01-03 for 1.000.000$ an equity-market ETF (SPY, DIA, QQQ or IWM). The hedge goes the 1st VIX-Future long. The position is rolled over 5 trading-days before maturity. The minimum-maturity of a new VIX-Future must be 15-trading days. This reduces the number of roll-overs. The trading cost is the minimum spread of 0.05 or 50$ per Future and trade. One holds no Futures if the risk-probability is below 0.95. If the risk-probability is above 0.95 one buys for 20% of the current value of the ETF position VIX-Futures. If the ETF has currently a value of 1.000.000$ one would invest a value of 200.000$. If the Futures price is 20.0, one would buy 10 Futures. One can only trade a full Future. The volume is rounded to the nearest integer-value. One does not adjust the Futures-volume as long as the risk-probability is above 0.95. If the probability falls below 0.95 the Futures are sold.

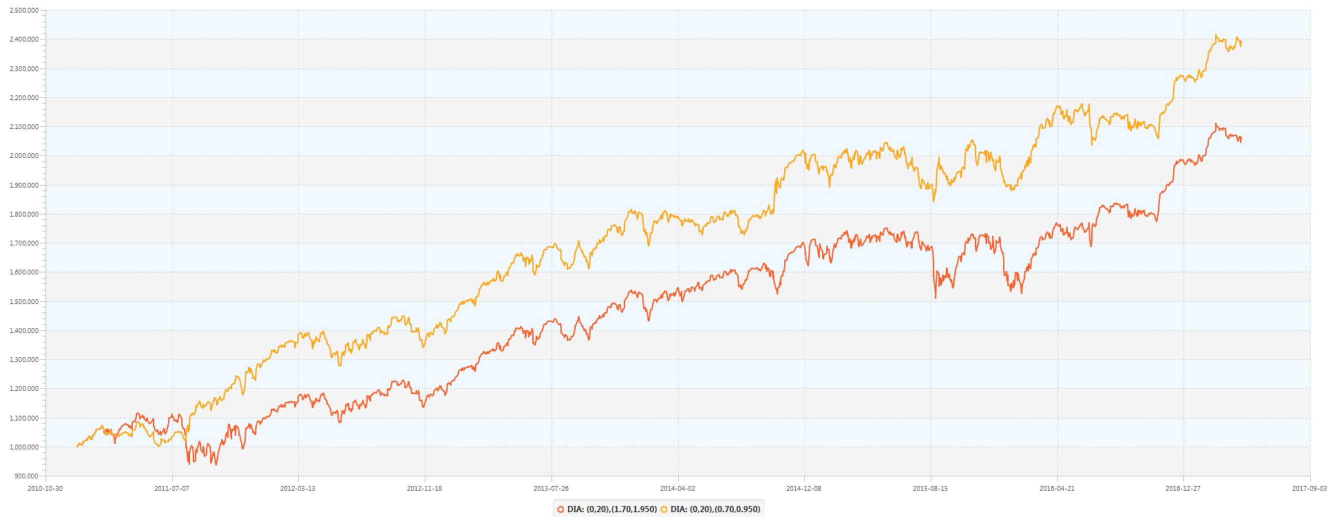
Note: The original paper used a slightly different trading strategy. One entered already at a risk of 0.7 a protection of 6% and the threshold for the full-protection was set to 0.90 with a protection of 18%. As there is a fast transition between the 2 states the results are similar.

The relative value of the Futures position can thus (considerably) differ from the initial allocation. The Futures are financed from a cash-account. But the cash-account is cleared at the end of each month. If the Futures made some profit, one increases the ETF allocation accordingly. If the cash balance is negative, one sells the same amount of the ETF. The ETF can also only be hold in full units. All the operations are up to rounding effects. The trading costs of the ETF are 2 Cent per trade. The overall transaction volume of the ETF is relative low. Different (higher) transactions costs change the results only marginally.

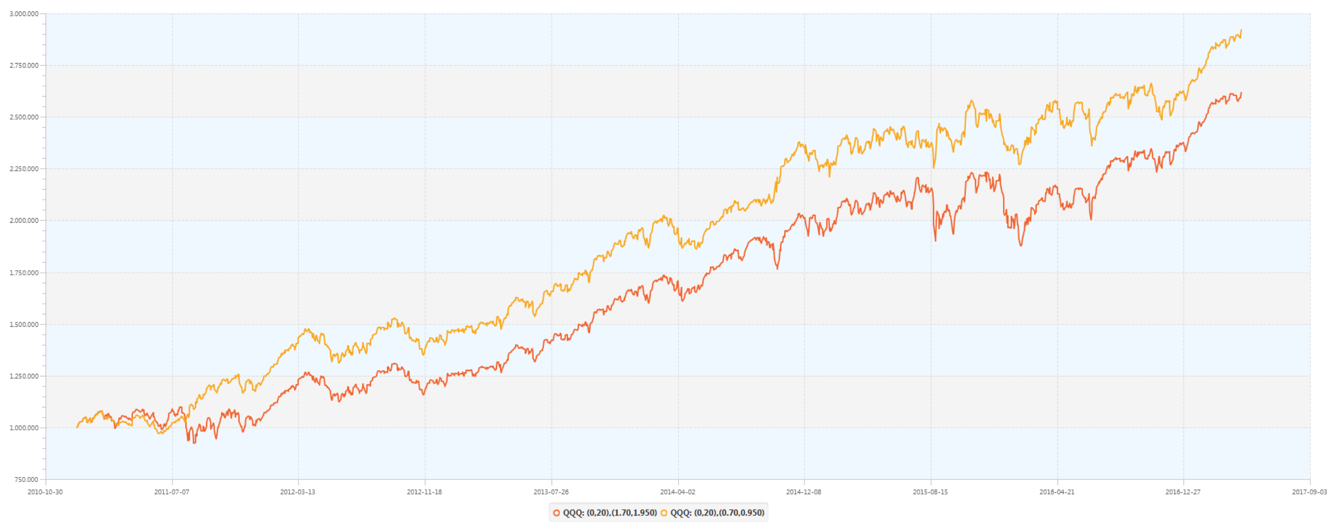
Graphic-2 shows in red the performance of the Buy&Hold of the SPY. The yellow chart is the performance of the Tail-Protection. The calculation uses the adjusted close-price. The same is done in Graphic-3 to 5 for the DIA, QQQ and IWM. For the detailed performance measures see Table-1 below.



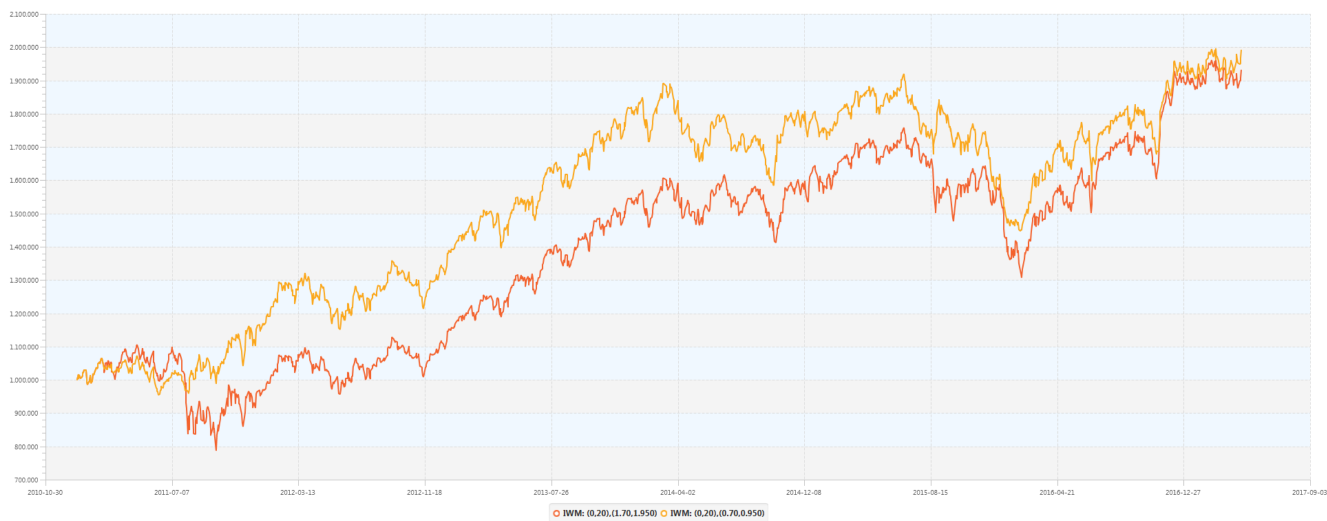
Graphic-2: SPY Buy&Hold (red) and HMM-Tail-Hedged (yellow) from 2011-01-03 till 2017-04-20



Graphic-3: DIA Buy&Hold (red) and HMM-Tail-Hedged (yellow) from 2011-01-03 till 2017-04-20



Graphic-4: QQQ Buy&Hold (red) and HMM-Tail-Hedged (yellow) from 2011-01-03 till 2017-04-20



Graphic-5: IWM Buy&Hold (red) and HMM-Tail-Hedged (yellow) from 2011-01-03 till 2017-04-20

T

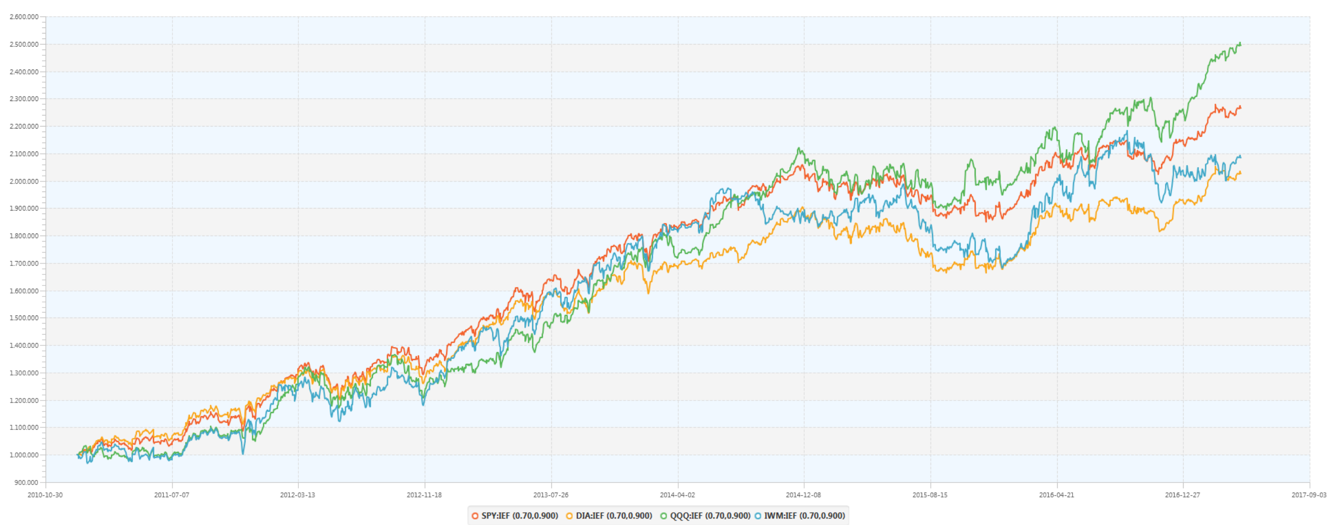
Strategy	Return	Sharpe	Drawdown
SPY-Buy&Hold	112.7	0.86	18.6
SPY-VIX-Futures	140.0	0.98	9.7
DIA-Buy&Hold	106.2	0.84	16.0
DIA-VIX-Futures	139.5	0.95	10.1
QQQ-Buy&Hold	161.6	0.90	16.1
QQQ-VIX-Futures	191.9	0.98	12.0
IWM-Buy&Hold	93.0	0.60	28.9
IWM-VIX-Futures	98.9	0.67	24.5

Table-1

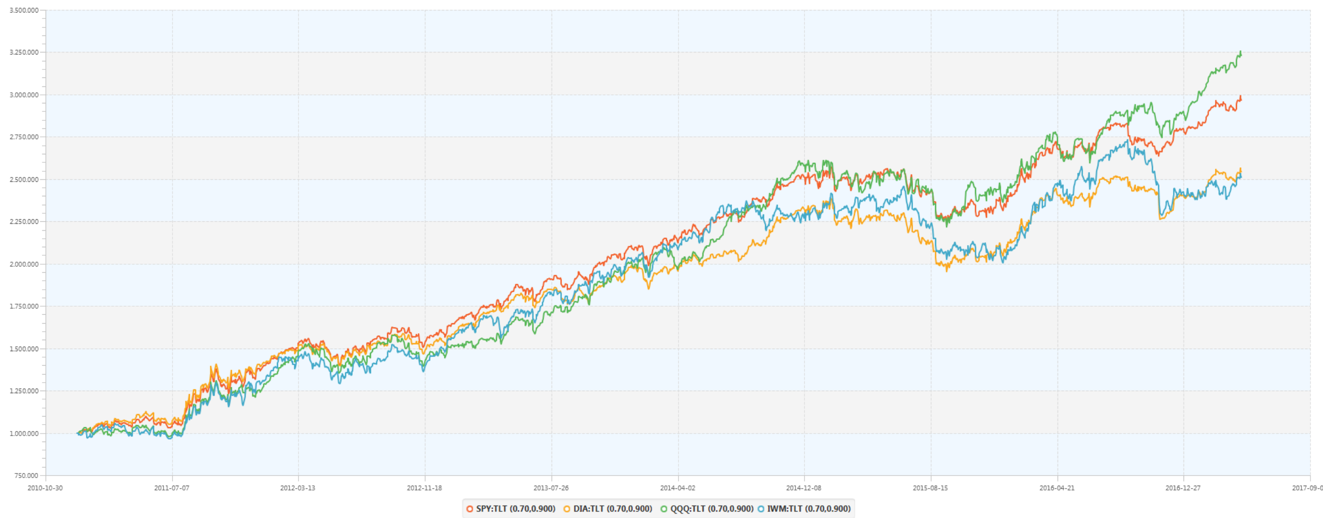
As can be seen in Table-1 the Tail-Protection improves for each ETF the Return, the Sharpe-Ratio and reduces the max. relative Drawdown. The effect is most pronounced for the SPY. This is to expect, because the VIX-Futures depend directly on the behavior the S&P-500. The improvement is similar for the DIA. The 2 indexes are highly correlated. It does not work so well for the IWM. The Russel-2000 under performed the S&P-500 from 2014 to 2016 by a wide margin. The VIX-Futures are in such periods no hedge. They just cost money. In August 2011 all the indexes crashed in almost the same way and the hedge worked for all ETFs quite fine.

Switching between Stocks and Treasuries:

An alternative to buying VIX-Futures is to switch in a bear-market to a treasury-ETF. Treasuries are considered as a save haven and gain value during market-turmoil. The effect is not as spectacular as for VIX-Futures, but treasuries are not systematically loosing value in a calm market environment.
Graphic-6 shows the performance of a strategy which sells above a risk-probability of 0.9 the equity-ETF and buys the IEF (7-10 years treasury index). If the risk-probability falls below 0.9 the IEF is sold and the old position is restored (the volume differs because the prices have changed). The calculations assume a trading-cost of 2 Cents per share and trade. The ETFs are very liquid and the bid-ask spread is currently 0.01 or 0.02.
Graphic-7 is the same strategy, but one switches to the TLT (20+ years treasuries). The TLT is considerable more volatile than the IEF. The overall return increases significantly, the risk-adjusted measures are for both treasuries about the same. The risk-adjusted performance is for the SPY, DIA and QQQ close to the VIX-Futures approach. It is a significant improvement for the IWM. The deviation of the IWM from the SPY is much less critical. For the detailed results see table-2 on the next page.



Graphic-6: SPY (red), DIA (yellow), QQQ (green) and IWM (blue) to IEF



Graphic-7: SPY (red), DIA (yellow), QQQ (green) and IWM (blue) to TLT

Strategy	Return	Sharpe	Drawdown
SPY:IEF	126.5	1.03	10.2
SPY:TLT	196.8	1.04	13.5
DIA:IEF	102.6	0.94	12.8
DIA:TLT	154.3	0.93	18.2
QQQ:IEF	149.5	0.96	11.7
QQQ:TLT	223.1	1.00	15.4
IWM:IEF	108.5	0.84	15.5
IWM:TLT	151.3	0.83	18.5

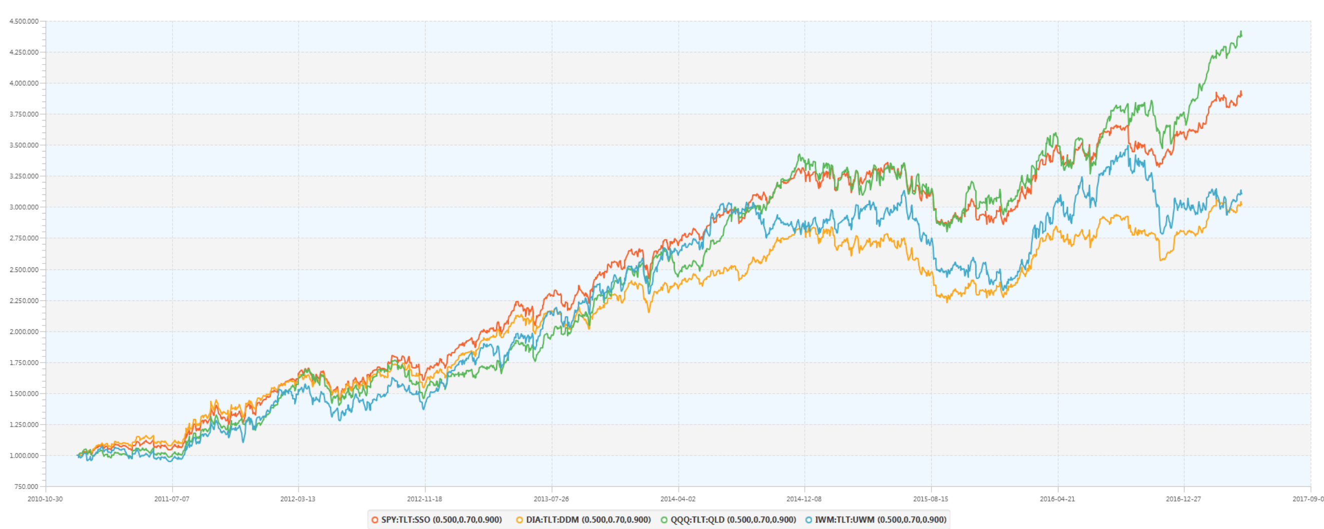
Table-2

Strategy	Return	Sharpe	Drawdown
SPY:SSO:TLT	290.04	1.02	16.42
DIA:DDM:TLT	201.96	0.89	22.05
QQQ:QLD:TLT	338.04	0.96	18.39
IWM:UWM:TLT	210.06	0.8	25.88

Table-3

Switching between the Base, the Ultra and Treasuries:

One can try to boost the performance of the equity-ETFs by leverage. If the risk-probability is below 0.7 the strategy is invested for 50% in the Ultra, for 50% in the base ETF. This ratio is not rebalanced as long as one stays within the same risk-regime. Between 0.7 and 0.9 only the base ETF is hold, above 0.9 one switches to the TLT.



Graphic-8: SPY/SSO (red), DIA/DDM (yellow), QQQ/QLD (green) and IWM/UWM (blue) to TLT

The performance is shown in Table-3 on the previous page. The overall return is considerable improved, but the risk-adjusted measures decrease in comparison to the base combination. This is the well known effect of the volatility drag. One can not increase the leverage arbitrarily.

Conclusion:

The extended analysis confirms the rather promising results of the original paper. But it should be noted that the additional equity-market indexes are highly correlated with the SPY. The model exploits the intrinsic relation between the SPX and the VIX-Futures Term-Structure. It can and should only work with highly correlated indexes. It is inline with the model assumptions that the strategy does not work as well for the IWM.

The only reasonable approach for testing the validity of the model would be a Monte-Carlo simulation. One would have first to have a realistic simulation model for the equity-index. This is due to volatility-clustering already a quite challenging task. Additionally one needs to model the relation between the SPX and the VIX-Futures Term-Structure. I have made several attempts to develop such an overall model. The results were not satisfactory. I also do not know of any other – practically useful – model.

Further Work:

No further work is planed for the time being.

Acknowledgment:

This rework and extension of the original HMM paper was inspired by discussions with Irene Miniböck, chairwoman of the Tiger Trading Club.

References:

1. Ch. Donninger: Timing the Tail-Risk-Protection of the SPY with VIX-Futures by a Hidden Markov Model.The Wool-Milk-Sow Strategy. Sibyl-Working-Paper, April 2017
2. William Press et. al: Numerical Recipes, 3rd Edition, Chapter 16.3.3
3. Andrew Fraser: Hidden Markov Models and Dynamical Systems. Chapter 3.1.2