Original Article

Detecting change points in VIX and S&P 500: A new approach to dynamic asset allocation

Peter Nystrup
is an industrial PhD student at Sampension, a Danish pension fund, and in the Department of Applied Mathematics and Computer Science at the Technical University of Denmark. He is a financial engineer and holds a Bachelor's and Master's degree in Engineering, Mathematical Modeling and Computation with honors from the Technical University of Denmark. His research focuses on detection of regime shifts in financial time series, methods for estimating statistical models with time-varying parameters, and development and implementation of asset allocation strategies using dynamic optimization methods.

Bo William Hansen
is the head of investment analysis at Sampension, where he, among other things, is responsible for the overall department's input to strategic asset allocation, risk management, ALM, performance analysis and Solvency II. Before joining Sampension in 2008, he worked with risk management and macroeconomic modeling at the Danish Central Bank, Nationalbanken. He holds a Master's degree in Economics from the University of Aarhus, Denmark and a Master of Commerce from The University of New South Wales, Australia.

Henrik Madsen
is a professor and head of the Section for Dynamical Systems in the Department of Applied Mathematics and Computer Science at the Technical University of Denmark. He earned his PhD in Statistics from the Technical University of Denmark in 1986. His main research interest is analysis and modeling of stochastic dynamical systems. This includes signal processing, time series analysis, identification, estimation, grey-box modeling, forecasting, optimization and control. The applications are mostly related to energy systems, informatics, environmental systems, bioinformatics, biostatistics, process modeling and finance. He has authored or co-authored approximately 500 papers and 12 books.

Erik Lindström
is an associate professor at the Centre for Mathematical Sciences at Lund University, Sweden. He earned his Master's degree in Engineering Physics in 2000 and his Master's degree in Business and Economics in 2001, followed by a PhD in Mathematical Statistics in 2004, all from Lund University. He has a great interest in teaching, being part of the university's Pedagogical Academy and being awarded the distinction 'Excellent Teaching Practitioner' in 2013. His research ranges from statistical methodology (primarily time series analysis in discrete and continuous time) to financial mathematics and problems related to energy markets, having published extensively in all those fields.

Correspondence: Peter Nystrup, DTU Compute, Asmussens Allé, Building 303B, 2800 Kgs. Lyngby, Denmark.
E-mail: pnys@dtu.dk

ABSTRACT The purpose of dynamic asset allocation (DAA) is to overcome the challenge that changing market conditions present to traditional strategic asset allocation by adjusting portfolio weights to take advantage of favorable conditions and reduce potential draw-downs. This article proposes a new approach to DAA that is based on detection of change points without fitting a model with a fixed number of regimes to the data, without estimating any parameters and without assuming a specific distribution of the data. It is examined whether DAA is most profitable when based on changes in the Chicago Board Options Exchange Volatility Index or change points detected in daily returns of the S&P 500 index. In an asset universe consisting of the S&P 500 index and cash, it is shown that a dynamic

strategy based on detected change points significantly improves the Sharpe ratio and reduces the drawdown risk when compared with a static, fixed-weight benchmark.

INTRODUCTION

The financial crisis of 2007–2008 resulted in large losses for most static portfolios, leading to increased interest in dynamic approaches to asset allocation. The market ructions during the beginning of 2016 are a recent example of how abruptly the behavior of financial markets can change. Although some changes may be transitory, the new behavior often persists for several months after a change (Ang and Timmermann, 2012). The mean, volatility and correlation patterns in stock returns, for example, changed dramatically at the start of, and persisted through the crisis of 2007–2008.

Changes in market dynamics present a big challenge to traditional strategic asset allocation (SAA) that seeks to develop static 'all-weather' portfolios that optimize efficiency across a range of scenarios. The purpose of dynamic asset allocation (DAA) is to take advantage of favorable market conditions and reduce potential drawdowns by adjusting portfolio weights as new information arrives (Sheikh and Sun, 2012). DAA is distinct from tactical asset allocation. While the latter relies on forecasting, DAA is based on reacting to changes in market conditions. The goal of DAA is not to predict change points or future market movements, but to identify when a regime shift has occurred and then benefit from persistence of equilibrium returns and volatilities.

Some approaches to DAA exploit the relationship between observed regimes in financial markets and the phases of the business cycle while other approaches are based solely on market data. Regime-switching models, such as the hidden Markov model (HMM), are a popular choice for modeling the hidden state of financial markets because they are able to reproduce stylized facts of financial returns, including volatility clustering and leptokurtosis (see, for example, Rydén *et al*, 1998; Nystrup *et al*, 2015b). Furthermore, the inferred states can often be linked to phases of the business cycle (see Guidolin and Timmermann, 2007). Several studies have shown the profitability of DAA strategies based on regime-switching models (see, for example, Guidolin and Timmermann, 2007; Bulla *et al*, 2011; Kritzman *et al*, 2012; Nystrup *et al*, 2015a).

A DAA strategy has two components: a method for detecting change points and a strategy for changing the portfolio when a change point has been detected. In case of a strategy based on a regime-switching model, the two components fuse as the asset allocation is typically fully determined by the inferred regime. The pivotal question then becomes, how many regimes are needed and whether the regimes can be assumed to be stationary. Assuming a fixed number of regimes (typically between two and four) based on economic motivations, as it has often been done in the literature (see Guidolin, 2011), is unlikely to be optimal.

The approach taken in this article is very different. It is based on detection of change points without fitting a model with a fixed number of regimes to the data, without estimating any parameters and without assuming a specific distribution of the data. The change points will not necessarily correspond to turning points in the business cycle. When a change point has been detected, the only knowledge of the new

regime is the observations encountered between the change point and the time of detection. The portfolio adjustment is determined independently of the change point detection.

This article presents a new approach to DAA. It examines whether DAA based on non-parametric change point detection can provide better long-term results when compared with static, fixed-weight portfolios. The focus will be on stock returns as portfolio risk is typically dominated by stock market risk. It will be examined whether to test for location, scale or more general distributional changes and whether the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) or past returns of the S&P 500 index yield the best proxy for the prevailing regime. Finally, it will be examined whether DAA is most profitable when based on changes in the VIX or change points detected in daily returns of the S&P 500 index. The VIX is considered because of its forward-looking nature that will be discussed in the next section before introducing change point detection and the empirical results in the following sections.

THE VIX AND S&P 500

The CBOE VIX, introduced by the CBOE in 1993, is an important benchmark for stock market volatility. The VIX estimates expected market volatility by averaging weighted prices of 30-calendar-day S&P 500 options over a range of strike prices. It considers a model-free estimator of implied volatility and, thus, does not depend on any particular option pricing framework.[1] The VIX essentially offers a market-determined, forward-looking estimate of 1-month stock market volatility and is regarded as an indicator of market stress; the higher the VIX, the greater the fear (Whaley, 2000).

There is a significantly negative relationship between changes in the VIX and contemporaneous returns of the S&P 500 index as depicted in Figure 1. The negative
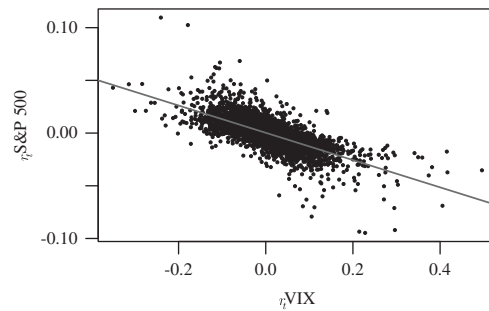

Figure 1: The leverage effect illustrated by contemporaneous log-returns of the VIX and the S&P 500 index.
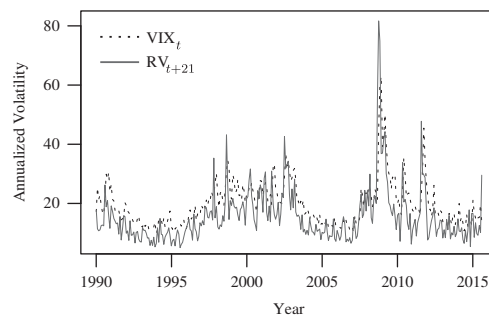

Figure 2: The VIX and the realized volatility of S&P 500 1-month ahead.
Note: Only one observation per month is shown.

correlation between stock returns and implied volatility captures the leverage effect first discussed by Black (1976). Racicot and Théoret (2016) found that the VIX embeds many properties of other macroeconomic and financial uncertainty measures, such as the growth of industrial production, the growth of consumer credit, long-term interest rates, term spreads and so on.

In Figure 2, the VIX is compared with the realized volatility of the S&P 500 index 1 month ahead.[2] It is evident that the VIX has had a persistent bias over realized volatility. The VIX is not a pure forecast, it is the price of volatility and, as such, includes a risk premium that varies over time. A negative risk premium for volatility partly explains the bias.

It appears from Figure 2 that spikes in realized volatility are followed by spikes in the VIX. Andersen *et al* (2007) showed that monthly volatility forecasts based on past realized volatility are more efficient than the VIX model-free implied volatility. It may,
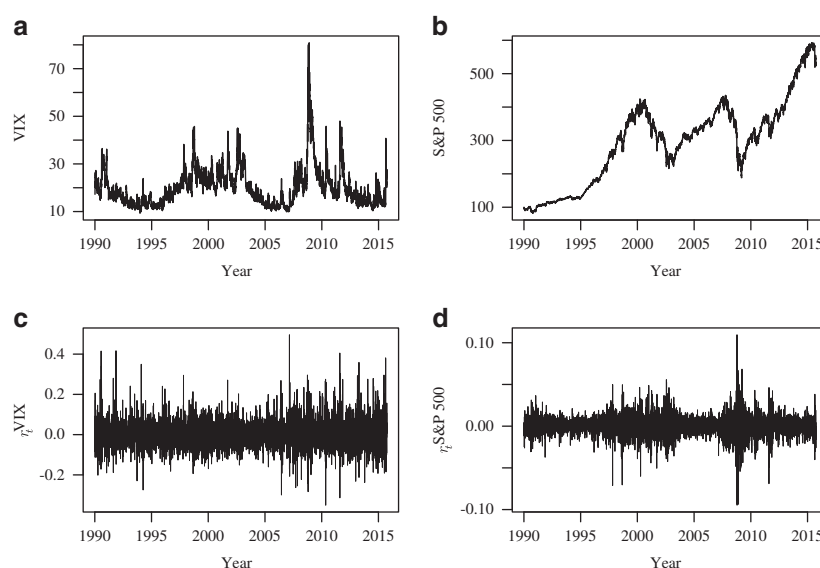
Figure 3: The VIX, the S&P 500 index and their daily log-returns.

however, still be the case that the VIX provides earlier signals about change points than past returns of the S&P 500 index.

The data analyzed is 6485 daily log–returns of the VIX and the S&P 500 index covering the period from January 1990 through September 2015. Figure 3 shows the two indices[3] together with their log-returns.[4] The volatility forms clusters as large movements tend to be followed by large movements and *vice versa*, as noted by Mandelbrot (1963).[5] This is most pronounced in the log-returns of the S&P 500 index, although it can be seen in the VIX series as well.

The volatility of the VIX appears to be higher when the level of the VIX is high. Persistence in the volatility of the VIX corresponds to persistence in the kurtosis of the S&P 500 returns, as the second moment of the VIX corresponds to the fourth moment of the S&P 500 returns. DAA aims to exploit this persistence of the volatility, as risk–adjusted returns, on average, are substantially lower during turbulent periods, irrespective of the source of turbulence, as shown by Kritzman and Li (2010).

The log-return series are characterized by a large number of exceptional observations.

Table 1: The first four moments of the log-returns

	VIX	S&P 500
Mean	0.000054	0.00026
Standard deviation	0.063	0.011
Skewness	0.69	−0.24
Kurtosis	7.2	11.7

Table 1 shows the first four moments of the log-returns of the two indices. The moments of the two return series are quite different, but the kurtosis is well above three for both series implying that the tails of the unconditional distributions are heavier than those of the Gaussian distribution. Assuming the data is Gaussian would cause occasional large values to be interpreted as change points, even though they should more correctly be classified as outliers, as discussed by Ross (2013).

CHANGE POINT DETECTION

A frequently used analogy for financial returns in the case of regime-switching models is that of a person's heart rate. While the person sleeps, a low average heart rate with low volatility is observed. When the person wakes up, there is a sudden rise in the heart rate's average level and its volatility.

Without actually seeing the person, it can reasonably be concluded based on observations of the heart rate whether he or she is awake or sleeping. Rather than distinguishing between a fixed number of states, such as awake or sleeping, the aim of change point detection is to detect when the distribution of the heart rate changes.

Change detection problems, where the goal is to monitor for distributional shifts in a sequence of time-ordered observations, arise in many diverse areas. They have been studied extensively within the field of statistical process control, where the goal is to monitor the quality characteristics of an industrial process in order to detect and diagnose faults.

The task is to detect whether a data sequence contains a change point. Using the framework of Ross *et al* (2011), if no change point exists, the observations are assumed to be identically distributed. If a change point exists at time τ, then the observations are distributed as:

$$X_i \sim \begin{cases} F^0 & \text{if } i < \tau, \\ F^1 & \text{if } i \geq \tau. \end{cases} \quad (1)$$

In other words, the variables are assumed iid with some distribution F^0 before the change point at $t = \tau$ and iid with a different distribution F^1 after. The location of the change point is unknown, and the problem is to detect it as soon as possible. The change point methodology can also be applied to sequences that are not iid between change points, by first modeling the data sequence in a way that yields iid one-step-ahead forecast residuals and then performing change detection on these (Ross *et al*, 2011).

Most traditional approaches to change detection assume that the distributional form of the data is known before and after the change with only the parameters being unknown. Classical methods for this problem include the CUSUM method (Page, 1954), exponentially weighted moving average (EWMA) charts (Roberts, 1959) and generalized likelihood ratio tests

(Siegmund and Venkatraman, 1995). These assumptions, however, rarely hold in sequential applications. Typically, there is no prior knowledge of the true distribution or assumptions made about the distribution may be incorrect (Ross *et al*, 2011).

A non-parametric approach

The non-parametric (distribution-free) change detection approach applied in this article is based on Ross *et al* (2011) and the implementation by Ross (2015). It does not assume that anything is known about the distribution of the data before monitoring begins; it is thus an example of a self-starting technique. The main advantage of the approach being self-starting is that it can be deployed out-of-the-box without the need to estimate parameters of the data distribution from a reference sample before monitoring.

In a typical setting, a sequence of observations $x_1, x_2, \ldots$ are received from the random variables $X_1, X_2, \ldots$. The distribution of X_i is given by (1), conditional on the change point τ that the task is to detect. Suppose t points from the sequence have been observed. For any fixed $k < t$, the hypothesis that a change point occurred at the kth observation can be written as:

$$H_0 : \forall i \, X_i \sim F_0, \qquad H_1 : X_i \sim \begin{cases} F_0 & \text{if } i < k, \\ F_1 & \text{if } i \geq k. \end{cases}$$

A two-sample hypothesis test can be used to test for a change point at k. Let $D_{k,t}$ be an appropriately chosen test statistic. For example, if the change is assumed to take the form of a shift in location and the data is assumed to be Gaussian, then $D_{k,t}$ will be the statistic associated with the usual t-test. If $D_{k,t} > h_{k,t}$ for some appropriately chosen threshold $h_{k,t}$, then a change is detected at location k. As no information is available concerning the location of the change point, $D_{k,t}$ has to be evaluated at all values of $1 < k < t$. If

$$D_{\max,t} = \max_k D_{k,t}, \quad (2)$$

then the hypothesis that no change has occurred before the tth observation is rejected if $D_{\max, t} > h_t$ for some threshold h_t. The estimate $\hat{\tau}$ of the change point is then the value of k that maximizes $D_{k, t}$. This formulation provides a general method for non-sequential change point detection on a fixed size data set. It can also be applied in the case where points are sequentially arriving over time, by repeatedly recomputing the test statistic as each new observation is received.

When the task is to detect a change in a data sequence where no information is available regarding the pre- or post-change distribution, the approach of Ross *et al* (2011) is to replace $D_{k, t}$ with a non-parametric two-sample test statistic that can detect arbitrary changes in a distribution. The algorithm would proceed as above, with this statistic evaluated at every time point, and the maximum value being compared with a threshold h_t. In many situations, a more powerful test can be found by restricting attention to the case when the pre-change distribution F_0 undergoes a change in either location or scale:

- Location Shift: $F_1(x) = F_0(x+\delta)$.
- Scale Shift: $F_1(x) = F_0(\delta x)$.

This corresponds to a change in either the mean or volatility of financial returns. Although it is slightly more restricted than testing for arbitrary changes, in practice any change in F_0 is likely to cause a shift in location or scale, and thus can be detected.

Non-parametric change detection is typically perceived to be less powerful than parametric change detection, but Ross *et al* (2011) showed that, although parametric Gaussian change detection using the t- and F-test outperforms the non-parametric methods when detecting larger-sized changes in the parameters of a Gaussian distribution, the difference in performance is not excessive, and the non-parametric tests actually outperform for smaller-sized changes. For heavy-tailed data the non-parametric methods significantly outperform the parametric Gaussian methods.

Of the non-parametric test statistics that were considered for this article, the tests that – in agreement with Ross *et al* (2011) – were found to be most powerful are the Mann–Whitney test for changes in location (Mann and Whitney, 1947), the Mood test for changes in scale (Mood, 1954) and the Lepage test for joint monitoring of changes in location and scale (Lepage, 1971). The Cramér–von Mises and Kolmogorov–Smirnov tests (see Ross and Adams, 2012) were also considered, but they were found to be slower at detecting distributional changes when compared with the Lepage test.

The mentioned tests use only the rank of the observations.[6] The Mood test (Mood, 1954), for example, is based on the observation that if $n = n_A + n_B$ points are spread over two samples A and B, then assuming no tied ranks, the expected rank of each point under the null hypothesis that both samples are identically distributed is $(n + 1)/2$. The Mood test uses a test statistic that measures the extent to which the rank of each observation deviates from the expected value:

$$M' = \sum_{x_i \in A} \left(r(x_i) - \frac{(n+1)}{2} \right)^2, \qquad (3)$$

where $r(x_i)$ denotes the rank of x_i in the pooled sample.

The distribution of the Mood statistic is independent of the distribution of the underlying random variables with mean and variance

$$\mu_{M'} = n_A \frac{(n^2 - 1)}{12},$$

$$\sigma^2_{M'} = n_A n_B (n+1) \frac{(n^2 - 4)}{180}.$$

Taking the absolute value of the standardized test statistic

$$M = \left| \frac{(M' - \mu_{M'})}{\sigma_{M'}} \right| \qquad (4)$$

ensures that both increases and decreases in scale can be detected.

EMPIRICAL RESULTS

The empirical testing shows that the best result is obtained by focusing on shifts in the scale parameter. It is not surprising that there is little information in the mean values (see, for example, Merton, 1980). Testing for more general distributional changes leads to approximately the same change points being identified with a larger detection delay compared with when focusing on shifts in scale. This finding applies to both the VIX and the S&P 500 return series.

In order to avoid a large number of false alarms, the expected time between false positive detections, also referred to as the average run length, is set to 10 000. The choice of average run length is a tradeoff between false positive detections and the delay in detecting changes. A total of 27 change points are detected in each of the two series when using the Mood test for changes in scale (Mood, 1954) as implemented by Ross (2015).

The detected change points are compared in Figure 4. Every second regime is shaded to make it easier to identify the change points. There is no further information in the shading. The detected change points seem very intuitive when shown together with the S&P 500 returns because of the distinct volatility clusters, while the change points detected in the log-returns of the VIX seem more intuitive when shown together with the index. The distance between the change points varies considerably from a few days up to 6 years. This is different from what would be expected if the change points were based on a business cycle indicator.

Although it is interesting in itself to analyze the change points, it is the times of detection that really matter. It is not possible to conclude whether changes in the VIX or past returns on the S&P 500 index yield the earliest signal about change points as not all
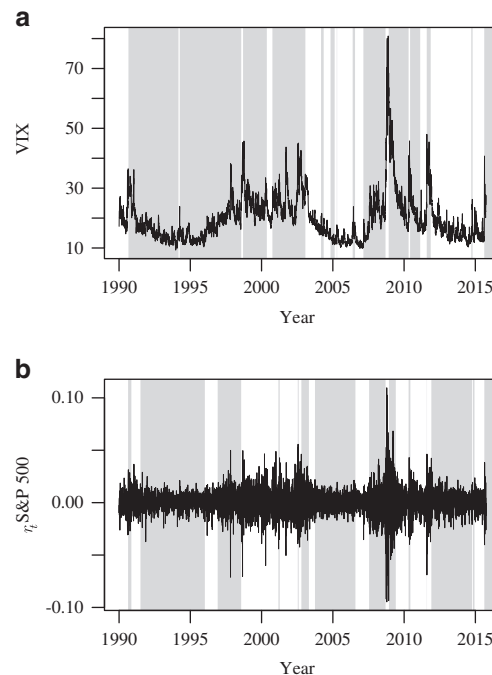

Figure 4: Comparison of the change points detected in the VIX and the S&P 500 log-returns.

the detected change points coincide. The next step is, therefore, to implement and test a DAA strategy based on the detected change points in order to determine which of the indices provides the most profitable signals.

Test procedure

The testing is done one day at a time in a live-sample setting to make it as realistic as possible. The first 21 observations, corresponding to 1 month, are used to determine the initial allocation. From that point onwards, it is tested each day whether a change point has occurred. If a change point is detected to have happened on day $\tau < t$ after the log-return of the index on day t has been added to the sample, then the observations from the time of the change point τ until the time of the detection t are used to estimate the volatility in the new regime. If, based on the new estimate of the volatility, the allocation to the stock index changes, then the new allocation is implemented at the closing of day $t + 1$, that is there is assumed to be a 1-day delay in the implementation. The asset allocation
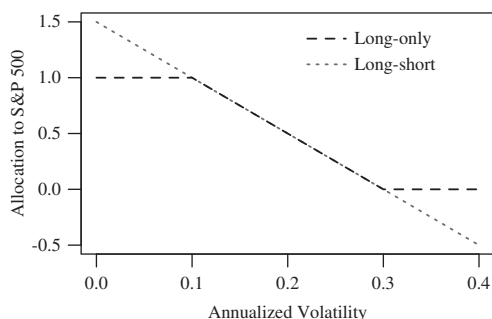
Figure 5: Two simple strategies where the allocation to S&P 500 is a function of the annualized volatility.

Table 2: The performance of two dynamic strategies based on the S&P 500 change points

Index/Strategy	AR	SD	SR	MDD
S&P 500	0.071	0.18	0.39	0.57
Long-only strategy	0.056	0.09	0.62	0.31
Static Portfolio I	0.047	0.11	0.43	0.39
Long–Short strategy	0.057	0.16	0.36	0.47

Note: This table summarizes the annualized return, standard deviation, Sharpe ratio and maximum drawdown for the S&P 500 index, the two dynamic strategies based on the S&P 500 change points, and a static portfolio that has the same average exposure (61 per cent) to the S&P 500 index as the Long-only strategy.

remains unchanged otherwise. The closing level on day $t + 1$ is then included in the sample and it is tested whether a new change point has occurred. The portfolio is not rebalanced until the next change point is detected.

Upon the detection of a change point, the volatility is estimated as the square root of an EWMA of the past variance

$$\text{EWMA}_t = \lambda \text{EWMA}_{t-1} + (1 - \lambda) r_t^2. \quad (5)$$

The forgetting parameter λ is set to 0.95, corresponding to an effective memory length of 20 trading days or roughly 1 month. This EWMA is found to be a better forecast of future, 1-month volatility than the VIX.

The tested strategies are fairly simple as the purpose is not to outline the optimal strategy, but rather to discuss the profitability of a DAA approach based on change point detection. Assuming the average long-term stock volatility is 20 per cent, a simple strategy is to allocate 50 per cent to stocks and 50 per cent to cash when the volatility is 20 per cent. If, instead, the volatility is 10 per cent, then the entire portfolio is allocated to stocks. If the volatility is 30 per cent, then the entire portfolio is allocated to cash. In Figure 5, this simple (linear) long-only strategy with no leverage is shown as the dashed line. The dotted line corresponds to the same allocation function with leverage and short-selling allowed. This strategy is referred to as Long–Short to distinguish it from the Long-only strategy. These strategies are likely not optimal, but attempting to optimize them may introduce a significant back-testing bias.

Performance of DAA strategies

S&P 500 change points

In Table 2, the performance of the two dynamic strategies based on the S&P 500 change points is compared with the performance of the S&P 500 index and a static, fixed-weight portfolio. The Static Portfolio I is rebalanced daily to have a fixed allocation of 61 per cent to the S&P 500 index that equals the average allocation of the Long-only strategy to the stock index over the period from February 1990 through September 2015. The remaining 39 per cent are allocated to cash that is assumed to yield zero interest.

The Long-only strategy has the highest Sharpe ratio (SR) with an annualized return (AR) of 5.6 per cent and an annualized standard deviation (SD) of 9 per cent.[7] Although the realized return is lower than for the S&P 500 index, the SR is significantly higher. The Long–Short strategy has the same AR as the Long-only strategy but a higher SD and maximum drawdown (MDD), and, consequently, a lower SR.

The Long-only strategy has a lower SD and MDD[8] than the Static Portfolio I (that has the same average exposure to the S&P 500 index), but the realized return is higher as long as transaction costs do not exceed 188 basis points per one-way transaction. This is when ignoring the costs associated with rebalancing to static weights, so the break-even transaction cost is higher than 188 basis points.
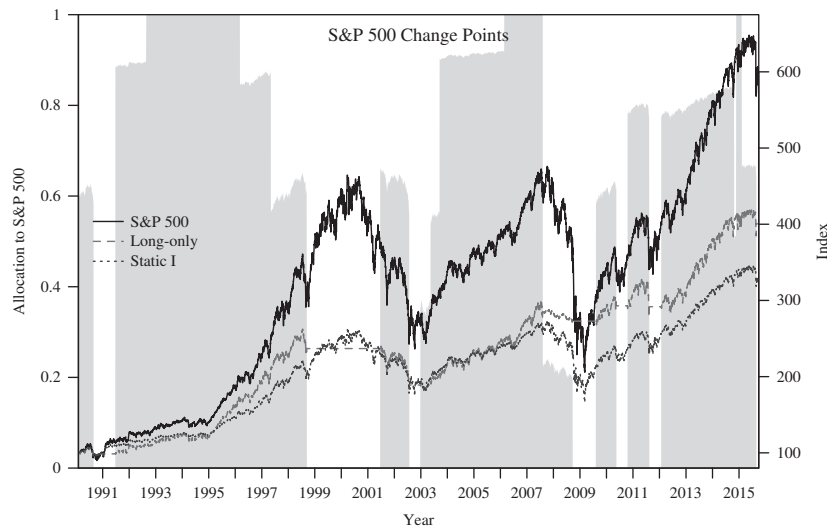
Figure 6: The development of the Long-only strategy based on the S&P 500 change points compared with a static, fixed-weight portfolio and the S&P 500 index (right axis). The shaded areas show the Long-only strategy's allocation to the S&P 500 index (left axis).
Note: The legends are sorted according to the final index value.

Figure 6 shows the development of the Long-only strategy, the Static Portfolio I and the S&P 500 index. The shaded areas show the Long-only strategy's allocation to the stock index. The times of detection of the change points that result in allocation changes are visible from the plot. On the basis of the allocation changes the volatility forecasts can be inferred. There appears to be many different levels of volatility. The choice of not rebalancing the dynamic portfolio between change points tilts it toward a momentum strategy. The effect, however, appears to be moderate. In case of a short position, thought of as implemented using a futures contract, the effect of not rebalancing is the opposite, as the size of the short position increases when the index goes up and decreases when the index goes down.

Most of the Long-only strategy's outperformance relative to the static, fixed-weight portfolio occurred during the financial crisis in 2008. It is clear why the Long–Short strategy is not profitable as the Long-only strategy is primarily fully allocated to cash around the peak in 2000 and on the way out of the crisis in 2008, when the market rebounded.

Table 3: The performance of two dynamic strategies based on the VIX change points

Index/Strategy	AR	SD	SR	MDD
S&P 500	0.071	0.18	0.39	0.57
Long-only strategy	0.062	0.10	0.64	0.24
Static Portfolio II	0.049	0.12	0.42	0.40
Long–Short strategy	0.059	0.15	0.40	0.37

Note: This table summarizes the annualized return, standard deviation, Sharpe ratio and maximum drawdown for the S&P 500 index, the two dynamic strategies based on the VIX change points, and a static portfolio that has the same average exposure (64 per cent) to the S&P 500 index as the Long-only strategy.

VIX change points

Table 3 summarizes the performance of the same strategies when based on change points detected by analyzing the daily changes in the VIX. The Static Portfolio II has a fixed allocation of 64 per cent to the S&P 500 index that equals the average allocation of the Long-only strategy to the stock index over the period.

The Long-only strategy has a higher AR and SR and a lower MDD than before. The AR is still lower than that of the index, but the difference is smaller. The AR of the Long-only strategy exceeds that of the Static Portfolio II as long as transaction costs do not
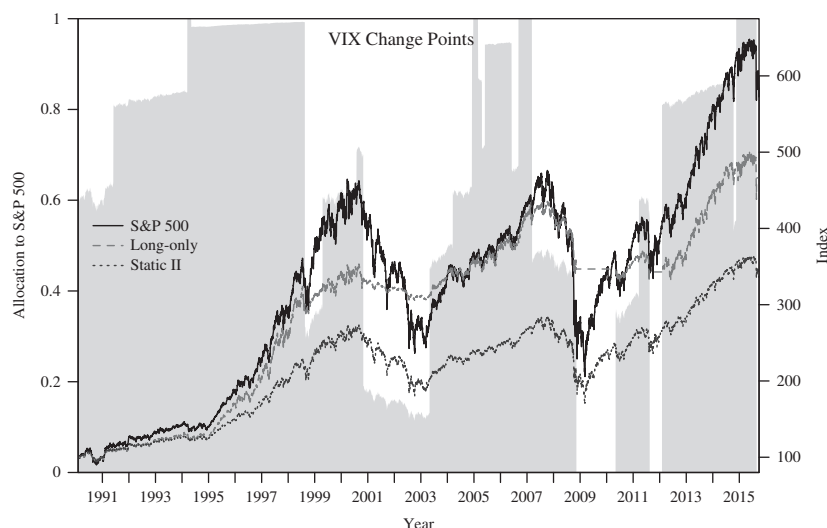
Figure 7: The development of the Long-only strategy based on the VIX change points compared with a static, fixed-weight portfolio and the S&P 500 index (right axis). The shaded areas show the Long-only strategy's allocation to the S&P 500 index (left axis).
Note: The legends are sorted according to the final index value.

exceed 372 basis points per one-way transaction.

The performance of the Long–Short strategy is similar to the index in terms of SR and it has a lower MDD. This is a significant improvement compared with when based on the S&P 500 change points, but it is still less profitable than the Long-only strategy.

Figure 7 shows the development of the Long-only strategy, the Static Portfolio II and the S&P 500 index. It was evident from Figure 4 that not all the change points are coinciding between the two series and by comparing Figure 6 and Figure 7 it appears that this it true for the detection times as well.

Compared with Figure 6, the Long-only strategy based on the VIX change points performs better in the years leading up to the burst of the dot-com bubble in year 2000. The allocation to S&P 500 is significantly reduced in 1998, around the same time as the allocation was reduced to zero when based on the S&P 500 change points, but the allocation is then increased to about 60 per cent in 1999. This allocation is retained all the way to the peak in year 2000. Toward the peak in 2008 the Long-only strategy performs worse when based on the VIX change points. The

allocation to S&P 500 is reduced for the first time already at the beginning of 2007, but it is not reduced all the way to zero until November 2008. In Figure 6, the allocation was reduced to zero in September 2008, only a few days after Lehman Brothers filed for bankruptcy.

The superior performance of the Long-only strategy compared with the Static Portfolio II increases steadily toward the peak in year 2000, through the subsequent downturn, and toward the peak in year 2008. The difference in performance is then reduced over the following 3 years during the market rebound before it again increases. The outperformance is built up gradually through the 25-year period and does not come from the financial crisis in 2008, on the contrary, the gap in performance compared with the static portfolio shrunk a lot in the wake of the crisis.

Performance of switching strategies

Although the tested strategies are fairly simple, they can be made even simpler by only allowing one asset in the portfolio at a time. Rather than having the allocation to the

S&P 500 index be a linear function of the forecasted volatility, the allocation is either 100 or 0 per cent in the long-only case and 100 or −100 per cent in the long–short case depending on whether the forecasted volatility is above or below 20 per cent. This threshold is chosen based on the assumption that the average long-term stock volatility is 20 per cent, which is an arbitrary choice. These simple switching strategies are similar to a regime-switching approach.

The testing of the switching strategies is carried out one day at a time in a live-sample setting exactly as before. The focus will be on the change points detected by analyzing the daily changes in the VIX as the DAA strategies already tested were most profitable when based on these. Past returns of the S&P 500 index are still used to estimate the volatility upon the detection of a change point.

Table 4 summarizes the performance of the switching strategies based on the VIX change points. The Static Portfolio III is rebalanced daily to have a fixed allocation of 66 per cent to the S&P 500 index, which equals the average allocation of the Long-only switching strategy to the stock index over the period.

The performance of the Long-only switching strategy is notable. Not only does it have a much higher SR and a significantly lower MDD than the S&P 500 index, it also has a higher realized return as long as transaction costs do not exceed 93 basis points

Table 4: The performance of two switching strategies based on the VIX change points

Index/Strategy	AR	SD	SR	MDD
S&P 500	0.071	0.18	0.39	0.57
Long-only switching strategy	0.075	0.11	0.68	0.20
Static Portfolio III	0.051	0.12	0.42	0.41
Long–Short switching strategy	0.074	0.15	0.50	0.44

Note: This table summarizes the annualized return, standard deviation, Sharpe ratio and maximum drawdown for the S&P 500 index, the two switching strategies based on the VIX change points, and a static portfolio that has the same average exposure (66 per cent) to the S&P 500 index as the Long-only switching strategy.

per one-way transaction. This is under the assumption that there is no interest on cash. The realized return of the Long-only switching strategy exceeds that of the Static Portfolio III as long as transaction costs do not exceed 627 basis points per one-way transaction and, at the same time, the MDD is less than half.

The Long–Short switching strategy is only rebalanced when the position changes from long to short or from short to long. This strategy also has a higher AR and a lower SD and MDD compared with the S&P 500 index. The SR of 0.50 is higher than that of the index, but it does not compare with that of the Long-only switching strategy.

Figure 8 shows the development of the Long-only switching strategy, the Static Portfolio III and the S&P 500 index. The allocation is 100 per cent whenever it was above 50 per cent in Figure 7 and 0 per cent whenever it was below 50 per cent.

Compared with the index, the switching strategy falls behind on the way toward the peak in year 2000, then it gets ahead during the subsequent downturn and retains its lead until the peak in year 2008 when the index catches up. The switching strategy gets far ahead during the crash in 2008 and the difference compared with the index is at its largest at the market trough in 2009. The lead almost diminishes during the turbulent second half of 2011, but the switching strategy retains a small lead all the way to the end of the sample. It is clear that if the big loss in 2008 was removed from the sample, the dynamic strategies would not outperform the index in terms of absolute return.

Compared with the static, fixed-weight portfolio, the Long-only switching strategy gradually extends its lead, with the exception of a few short periods, from the beginning of the sample to the market trough in 2009. The following 3 years the static portfolio regains some of the loss before the difference is again extended from 2012 to the end of the sample. The Long-only switching strategy is far in front of the static portfolio throughout the
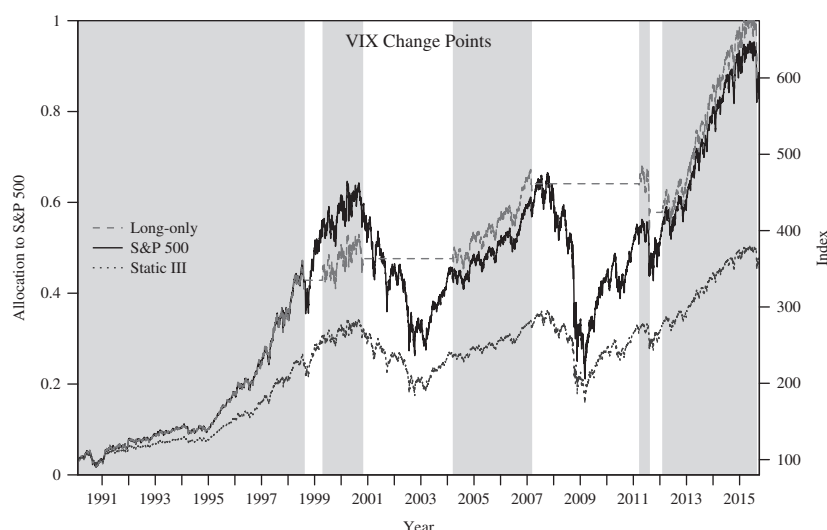
Figure 8: The development of the Long-only switching strategy based on the VIX change points compared with a static, fixed-weight portfolio and the S&P 500 index (right axis). The shaded areas show the Long-only strategy's allocation to the S&P 500 index (left axis).
Note: The legends are sorted according to the final index value.

sample as evidenced by the break–even transaction cost of 627 basis points.

It is easy to get the impression from Figure 8 that the applied change point method is faster at detecting increases in the volatility than decreases, as the switching strategy is much better at timing the downturns than the rebounds. The same conclusion was reached in Nystrup *et al* (2015a) based on regimes inferred with an HMM with time–varying parameters. From a comparison with Figure 7, it appears that the volatility remains high during the beginning of the rebound as the allocation is increased (to a level below 50 per cent) sooner than it can be seen from Figure 8. It is expected that gradual drift is harder to detect than abrupt changes.

Trading the VIX

Given the strong performance of the Long-only switching strategy based on the VIX change points, it is natural to wonder whether the performance can be replicated and possibly improved by trading the VIX itself. Instead of buying the S&P 500 when a change point is detected and the volatility is estimated to be below 20 per cent, the replicating

Table 5: The performance of two switching strategies based on the VIX change points

Index/Strategy	AR	SD	SR	MDD
S&P 500	0.044	0.21	0.21	0.57
Long-only switching strategy	0.045	0.09	0.47	0.18
VIX short-term futures inverse ER	0.117	0.64	0.18	0.92
Switching futures strategy	0.179	0.31	0.59	0.40

Note: This table summarizes the annualized return, standard deviation, Sharpe ratio and maximum drawdown for the S&P 500 index, the Long-only switching strategy, the S&P 500 VIX short-term futures inverse daily excess return index, and the switching futures strategy over the period from 20 December 2005 through 30 September 2015.

strategy would be to sell short–term VIX futures (see Ilmanen, 2012; Whaley, 2013; Simon and Campasano, 2014).

As long as the volatility remains low, a short position will accrue a roll yield as the futures price converges to the spot price, reflecting the gap between realized and implied volatility.[9] When a new change point is detected and the volatility is estimated to be above 20 per cent, the short position is terminated.

In Table 5, the performance of the switching futures strategy is compared with the S&P 500 VIX short-term futures inverse daily excess return index, the Long-only

switching strategy and the S&P 500 index over the period from 20 December 2005 through 30 September 2015.[10]

The return from selling short-term VIX futures is more than three times as high as the return of the S&P 500 index and the Long-only strategy that switches in and out of S&P 500, but the strategy is also more volatile.

The SR of the futures strategy is higher than that of the Long-only switching strategy when transaction costs are ignored. After accounting for transaction costs, the difference is most likely small, since the futures strategy involves more trading from rolling the short position (Whaley, 2013).

SUMMARY AND DISCUSSION

The new approach to DAA presented in this article is based on sequential hypothesis testing to detect change points without fitting a model with a fixed number of regimes to the data, without estimating any parameters and without making any assumptions about the distribution of the data. This approach is very robust, given that it is not based on a model or any assumptions about the data that can change going forward. It is a useful tool for dividing financial time series into regimes without making any assumptions about the number of regimes or the distribution of the data within the regimes.

Daily returns of both the VIX and the S&P 500 index were considered as input to the change point detection. The VIX was considered because of its forward-looking nature. It was not possible to conclude which of the two series that provided the earliest warning about change points as not all the detected change points were coinciding. The testing did show, however, that DAA based on the VIX change points was most profitable.

Simple switching strategies performed better than strategies where the allocation to the stock index was a linear function of the estimated volatility, despite the fact that the volatility assumed many different levels in the detected regimes.

The best performing strategy was a switching strategy that was fully invested in the S&P 500 index in the low-volatility state and cash in the high-volatility state. This strategy outperformed both the S&P 500 index and a static portfolio with the same average allocation to the stock index both in terms of SR and realized return, and it had a significantly lower tail risk. Owing to the assumption of zero interest on cash positions, there was no other source of performance than the index. A similar SR could be obtained by selling short-term VIX futures instead of buying the S&P 500 index in the low-volatility state.

The analysis focused on the S&P 500 price index because of its link to the VIX. If the price index was replaced by the total return version of the S&P 500 index and the return on cash was assumed to be the daily risk-free rate rather than zero, then the realized returns would be higher, but the relative performance of the strategies and the break-even transaction costs would be almost the same.

The tested strategies may be based on larger changes in allocation than most investors are willing to and/or allowed to implement. The excess return that can be obtained will simply be proportional to the fraction of the portfolio that is allocated to the dynamic strategy.

The presented results have important implications for portfolio managers with a medium to long-term investment horizon. Even without any level of forecasting skill it is not optimal to hold a static, fixed-weight portfolio. This new robust approach to DAA has the potential to improve (risk-adjusted) returns and reduce tail risk compared with traditional static SAA.

ACKNOWLEDGEMENTS

NOTES

1. The details of the computations are available at www.cboe.com/micro/vix/vixwhite.pdf.
2. The realized volatility is calculated, in accordance with the methodology used by S&P Dow Jones Indices, as $RV_t = 100 \times \sqrt{12 \sum_{i=t-20}^{t} log(P_i/P_{i-1})^2}$, where P_t is the closing price of the index on day t and log is the natural logarithm.
3. The S&P 500 index has been rescaled to start at 100 on 2 January 1990.
4. The log-returns are calculated using $r_t = log(P_t/P_{t-1})$, where P_t is the closing price of the index on day t and log is the natural logarithm.
5. A quantitative manifestation of this fact is that while returns themselves are uncorrelated, absolute and squared returns display a positive, significant and slowly decaying autocorrelation function.
6. The rank of the ith observation at time t is defined as $r(x_i) = \sum_{i \neq j}^{t} I(x_i \geq x_j)$, where I is the indicator function.
7. Neither of the portfolio return series display enough autocorrelation that it is necessary to adjust the annualized SDs.
8. The maximum drawdown is the largest relative decline from a historical peak in the index value.
9. The VIX futures curve has historically been in contango most of the time, especially when the VIX was below 20, as documented by Simon and Campasano (2014).
10. This is the longest history available for the S&P 500 VIX short-term futures inverse daily excess return index (SPVXSPI) provided by S&P Dow Jones Indices.

REFERENCES

Andersen, T.G., Frederiksen, P.H. and Staal, A.D. (2007) The Information Content of Realized Volatility Forecasts. Working paper. Illinois: Northwestern University.

Ang, A. and Timmermann, A. (2012) Regime changes and financial markets. *Annual Review of Financial Economics* 4(1): 313–337.

Black, F. (1976) Studies of stock price volatility changes. In *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*. Washington DC: American Statistical Association, pp. 177–181.

Bulla, J., Mergner, S., Bulla, I., Sesboüé, A. and Chesneau, C. (2011) Markov-switching asset allocation: Do profitable strategies exist? *Journal of Asset Management* 12(5): 310–321.

Guidolin, M. (2011) *Markov Switching Models in Empirical Finance*. Chapter 6. UK: Emerald Group Publishing, pp. 1–86.

Guidolin, M. and Timmermann, A. (2007) Asset allocation under multivariate regime switching. *Journal of Economic Dynamics & Control* 31(11): 3503–3544.

Ilmanen, A. (2012) Do financial markets reward buying or selling insurance and lottery tickets? *Financial Analysts Journal* 68(5): 26–36.

Kritzman, M. and Li, Y. (2010) Skulls, financial turbulence, and risk management. *Financial Analysts Journal* 66(5): 30–41.

Kritzman, M., Page, S. and Turkington, D. (2012) Regime shifts: Implications for dynamic strategies. *Financial Analysts Journal* 68(3): 22–39.

Lepage, Y. (1971) A combination of Wilcoxon's and Ansari-Bradley's statistics. *Biometrika* 58(1): 213–217.

Mandelbrot, B. (1963) The variation of certain speculative prices. *The Journal of Business* 36(4): 394–419.

Mann, H.B. and Whitney, D.R. (1947) On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* 18(1): 50–60.

Merton, R.C. (1980) On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8(4): 323–361.

Mood, A.M. (1954) On the asymptotic efficiency of certain nonparametric two-sample tests. *The Annals of Mathematical Statistics* 25(3): 514–522.

Nystrup, P., Hansen, B.W., Madsen, H. and Lindström, E. (2015a) Regime-based versus static asset allocation: Letting the data speak. *The Journal of Portfolio Management* 42(1): 103–109.

Nystrup, P., Madsen, H. and Lindström, E. (2015b) Long memory of financial time series and hidden Markov models with time-varying parameters. In *Proceedings of the 22nd International Forecasting Financial Markets Conference*, France: Forecasting Financial Markets Association.

Page, E.S. (1954) Continuous inspection schemes. *Biometrika* 41(1–2): 100–115.

Racicot, F.E. and Théoret, R. (2016) Macroeconomic shocks, forward looking dynamics and the behavior of hedge funds. *Journal of Banking and Finance* 62: 41–61.

Roberts, S.W. (1959) Control chart tests based on geometric moving averages. *Technometrics* 1(3): 239–250.

Ross, G.J. (2013) Modelling financial volatility in the presence of abrupt changes. *Physica A* 392(2): 350–360.

Ross, G.J. (2015) Parametric and nonparametric sequential change detection in R: The cpm package. *Journal of Statistical Software* 66(3): 1–20.

Ross, G.J. and Adams, N.M. (2012) Two nonparametric control charts for detecting arbitrary distribution changes. *Journal of Quality Technology* 44(2): 102–116.

Ross, G.J., Tasoulis, D.K. and Adams, N.M. (2011) Nonparametric monitoring of data streams for changes in location and scale. *Technometrics* 53(4): 379–389.

Rydén, T., Teräsvirta, T. and Åsbrink, S. (1998) Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics* 13(3): 217–244.

Sheikh, A.Z. and Sun, J. (2012) Regime change: Implications of macroeconomic shifts on asset class and portfolio performance. *The Journal of Investing* 21(3): 36–54.

Siegmund, D. and Venkatraman, E.S. (1995) Using the generalized likelihood ratio statistic for sequential detection of a change-point. *The Annals of Statistics* 23(1): 255–271.

Simon, D.P. and Campasano, J. (2014) The VIX futures basis: Evidence and trading strategies. *The Journal of Derivatives* 21(3): 54–69.

Whaley, R.E. (2000) The investor fear gauge. *The Journal of Portfolio Management* 26(3): 12–17.

Whaley, R.E. (2013) Trading volatility: At what cost? *The Journal of Portfolio Management* 40(1): 95–108.