Machine learning, economic regimes and portfolio optimisation

John M. Mulvey*, Han Hao and Nongchao Li

Bendheim Center for Finance,
Operations Research and Financial Engineering Department,
Princeton University, USA
Email: mulvey@princeton.edu
Email: hhao@princeton.edu
Email: nongchao@princeton.edu
*Corresponding author

Abstract: In portfolio models, the depiction of future outcomes depends upon a representative accounting of economic conditions. There is much evidence that crash periods display much different patterns than normal markets, suggesting that forecasting models ought to be based on multiple regimes. We apply two techniques from machine learning in our empirical study to improve robustness: 1) trend-filtering – to distinguish regimes possessing relatively homogeneous patterns; 2) a shrinkage/cross validation approach within a factor analysis of performance. A scenario-based portfolio model is proposed and designed to address multiple regimes. The worst-case events are well described within the framework, as compared with mean-variance Markowitz models that treat equally all historical performance.

Keywords: portfolio models; asset allocation; economic regimes; machine learning; classification; factor investing; worst-case events.

Reference to this paper should be made as follows: Mulvey, J.M., Hao, H. and Li, N. (2018) 'Machine learning, economic regimes and portfolio optimisation', *Int. J. Financial Engineering and Risk Management*, Vol. 2, No. 4, pp.260–282.

Biographical notes: John M. Mulvey is a Professor from the Department of Operations Research & Financial Engineering at Princeton University.

Han Hao is a PhD student in the Department of Operations Research & Financial Engineering at Princeton University. His research interests are portfolio optimisation, quantitative asset and risk management.

Nongchao Li is a PhD student in the Department of Operations Research & Financial Engineering at Princeton University. His research interests are quantitative investing and risk management.

1 Introduction

The success of an investment process depends upon several factors, including the accuracy of the forecasts, the timing of capital inflows and outflows, the need for capital in future time periods, the ability to act quickly when conditions change and so on.

For long-term investors, it is critical to reduce losses during sharp drawdown periods so that the investor's liabilities and goals will be met with a reasonable probability of success. To do so, a portfolio model must be able to depict future events with an eye to the worst-case, for example, maximum probability loss or maximum drawdown. However, traditional portfolio models may provide an inadequate representation of the left tail of the return distribution. Historical asset performance exhibits large negative skewness (excess kurtosis), for example and volatility spikes during crash periods.

It is well accepted that crash periods in economic markets are driven by high volatility and an extreme level of correlation – approaching unity in certain environments (Kyle and Xiong, 2001). What causes this condition? First, investor panic is one of the symptoms, especially by novices and individuals who are unfamiliar with a sharp drop. The situation is similar to a seasoned airline traveller versus a first-time travel when experiencing a period of high turbulence in a modern jet airline. Second, as evident in asset pricing models, one of the critical components is the overall level of risk premium. Crash periods are largely driven by systemic risks. Another contributing factor is the lack of liquidity, whereby markets do not easily clear for both buyers and sellers. Last, in certain cases, even riskless assets may be difficult to access. Recall the money market and commercial paper freeze in September 2008, when a number of money market funds experienced 30% to 60% withdrawal rates. These elements of a crash are commonly caused by 'neglected risks' (Gennaioli et al., 2012).

Empirical evidence indicates that the most asset categories exhibit excess negative skewness and excess kurtosis. Thus, the probability of a sharp loss is much higher than indicated by multinormal distributions. The 22% plunge in the S&P 500 on October 19, 1987 would be roughly 22 standard deviations below the mean return. The probability of a drop of this magnitude approaches zero in traditional portfolio models. There are many similar occurrences of non-normal performance. Take the early October 2011 period for the S&P 500 which experienced four consecutive days of 4% moves (down, up, down, up). Or the late 2008, subprime meltdown. Another characteristic of crash periods involves changing correlations and higher volatilities. In most crashes, volatility spikes (recall the October 2011 week) and correlations approach unity. There is clearly contagion in markets during crashes.

There are several advantages for employing multiple regimes in a scenario projection system. First, the worst-case environments can be modelled and parameterised directly with regard to the high volatility and high correlation. This step allows for more accuracy than by traditional static multi-normal assumptions. Second, the notion that crash periods have differing patterns than normal periods is well understood. Most investors who have experienced a major market drop will mention the panic that overwhelms the market. Comments include: there is no place to hide; I cannot unwind positions without massive market impact costs; the bid-ask spread and implied volatility grows to exceptional levels; and so on.

Third, empirical evidence is consistent with the contagion environment during crashes. The characteristics of crashes are relatively common: high volatility, high correlation and the presence of the flight-to-quality tactic to protect the investor's capital. While it is difficult to predict the timing of a crash, a carefully designed risk management system should be able to simulate the crash conditions to a reasonable degree and especially evaluate tactics and asset allocation rules that are designed to protect the investor's capital to a designated level. Investors should be prepared to accept the anticipated levels of drawdown, protection and the costs therein.

In the next sections, we apply a popular method in machine learning – trend-filtering – to the task of defining two regimes. First, we apply the trend-filtering approach to the S&P 500 index in order to identify two distinct regimes: labelled growth (normal) and contraction (crash). We compare the actual distribution of stock and other asset performance with the single regime d vis-à-vis the two-regime scenario generators with regard to various performance statistics. There are insights to be gained. The two-regime model better represents the historical patterns of performance than a single regime approach, especially with regard to the worst-case events. Second, the analysis illustrates the role of assets that have performed reasonably well under crash periods. These assets can be difficult to pinpoint under a traditional single-regime portfolio system. Also, for long-term investors, a multiple regime projection system displays some degree of mean reversion, which is consistent with historical evidence.

As a second study, we take up the issue of understanding the role of inflationary expectations in explaining asset performance. Again, we apply the trend-filtering algorithms, but instead of employing the S&P 500 index as the numeraire, we separate on an index of inflationary expectations. Again, we look at the performance of asset categories under the two regimes – in this case, rising inflationary expectations and lowering inflationary expectations. This analysis is meant to be explanatory, rather than as a strict forecasting tool.

In the final section, we construct an illustrative scenario portfolio model based on multiple regimes. The model focuses on worst-case risk measures and thereby aims to protect the investor's capital during contraction periods.

2 Regime identification via trend-filtering

A regime analysis is significant in asset allocation and asset-liability management for long-term investors because of the contagion and related effects during crash periods: the correlation between risky assets and volatility will greatly increase during the crash periods, thus creating severe difficulty in risk management and protecting investor capital with traditional portfolio models.

Many approaches have been proposed to identify differing economic regimes. These include hidden Markov models, structural breaks, advanced econometric models and more recently, social network analysis. See Ang and Bekaert (2002), Bae et al. (2013), Graflund and Nilsson (2003), Guidolin and Timmermann (2007, 2008), Ilmanen et al. (2014) and Mulvey and Reus (2016). It is outside the scope of this study to take up the interesting research topic of comparing approaches. Many existing methods are based on econometric models which assume a fixed structural model (mostly linear). However, financial return or macroeconomic data tends to be noisy and affected by myriad of factors. The state-of-the-art approach discussed in Mulvey and Liu (2016), trend-filtering, is non-parametric, data-driven and model-free. The algorithm was first introduced by Kim et al. (2009) and generalised in Tibshirani (2014). It aims to solve a 'filtering' problem, which separates trend from the input data and minimises estimation error. A general filtering problem for time series data can be written as

$$x_i = f\left(t_i\right) + \epsilon_i \tag{2.1}$$

where t_i is time and $f(t_i)$ shows a degree of trend.

In the trend-filtering algorithm, we manage to find some point correspondence between x and t. In specific, we construct a 'fitted' time series, β_i, that serves as the signal of the trend. This new time series can be obtained by solving the following optimisation problem:

$$\hat{\beta} - \underset{\beta \in \mathbb{R}^n}{\arg\min} \left\| x - \beta \right\|_2^2 + \lambda \left\| D^{(k+1)} \beta \right\|_1 \tag{2.2}$$

where $x = (x_1, \ldots, x_n)^T$, $\beta = (\beta_1, \ldots, \beta_n)^T$ are two vectors, $\lambda > 0$ is the hyper-parameter of this model and $D^{(k+1)}$ is the discrete difference operator. We can adjust k to obtain the desired shape of our fitted time series. In the basic form, we have $k = 0$ and the matrix is

$$D^{(1)} = \begin{bmatrix} -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}. \tag{2.3}$$

The output of the algorithm will be a piecewise constant time series which can be conveniently identified as a 'regime', for example, labelling positive periods as 'normal' and negative periods as 'crash'. The basic idea is to filter out the noise of the underlying time series and extract the trend. In the formula, the quantity we wish to find, β, is a n-dimensional vector which minimises a quantity of two parts, whereas the first part is the sum of square error of this vector and the input return data and the second part is the penalty term which penalises different signs of two consecutive data points. The second part will enable the algorithm to ignore a sudden change in the time series that likely to be triggered by noise, resulting in an indicator of the current momentum of the return series.

The trend-filtering algorithm can also be generalised so that the fitted time series is a piecewise polynomial. Consider d_k as the coefficients of $(-a + b)^{k+1}$, (e.g., d^2 is defined as $[-1, 3, -3, 1]$). Then, we define the $(k + 1)^{\text{th}}$-order difference operator

$$D^{(k+1)} = \begin{bmatrix} & d_k & & 0 & 0 & \ldots & 0 & 0 \\ 0 & & d_k & & 0 & \ldots & 0 & 0 \\ \vdots & & & & & & & \\ 0 & \ldots & \ldots & 0 & \ldots & & d_k & \end{bmatrix} \in \mathbb{R}^{(n-k-1)\times n} \tag{2.4}$$

and solve the optimisation problem (2.2).

We will briefly discuss below some properties of the trend-filtering estimator β. Detailed mathematical derivation is shown in Kim et al. (2009).

1 The estimator converges to x as $\lambda \to 0$.

2 The estimator converges to the best constant fit (i.e., the estimator is constant throughout the whole time) as $\lambda \to \infty$. In fact, the characteristic of the LASSO estimator enables the convergence as $\lambda \geq \lambda_{\max}$ for some $\lambda_{\max}$ large enough.

3 The estimator is piecewise constant for $k = 0$. To observe this, notice that our optimisation problem can be rewritten as a LASSO estimation:

$$\hat{\beta} = \underset{\theta \in \mathbb{R}^n}{\arg\min} \|x - H\theta\|_2^2 + \lambda \sum_{i=2}^{n} |\theta_i| \tag{2.5}$$

where $\theta = (\theta_1, \ldots, \theta_n)$ is the decision variable and H satisfies $H\theta = \beta$. Here, H is a lower triangular matrix that serves as the 'inverse' of D (but not exactly the inverse since D is not a square matrix). By the properties of LASSO estimator, the solution θ will likely to be sparse. Therefore, for many of the j, we will have $\beta_j = \beta_{j+1}$. The estimator is thus piecewise constant and we call the points *i kink points* if $\beta_i \neq \beta_{i+1}$.

4 The estimator is also piecewise constant, as a function of the hyper-parameter λ for $k = 0$. Although this hyper-parameter cannot be adjusted during the algorithm, we can use methods such as cross-validation to decide the optimal value that balances between bias and variance.

5 Similar to (3) and (4), the estimator will be piecewise k^{th} polynomial for any integer value of k. We show this for $k = 1$. By the property of LASSO estimator (Tibshirani, 1996), for a couple of j we will have $\beta_j - 2\beta_{j+1} + \beta_{j+2} = 0$, i.e., $\beta_j - \beta_{j+1} = \beta_{j+1} - \beta_{j+2}$. This equation shows that the fitted time series will be piecewise linear. In practice, higher values of k are rarely used because they may overfit.

Therefore, the trend-filtering algorithm provides flexibility in the form of input data. For series that we would like to separate high levels from low levels, for example, macroeconomic factors such as GDP or CPI, it is natural to optimise with the first-order difference matrix (2.3). The 'filtered' series will be a piecewise constant indicator of the macroeconomic conditions of each time period marked by the kink points. On the other hand, if the input is an index (in the next part we will see the result for S&P 500 index series), we can use the second-order ($k = 1$) difference matrix

$$D^{(2)} = \begin{bmatrix} 1 & -2 & 1 & \ldots & 0 & 0 & 0 \\ 0 & 1 & -2 & \ldots & 0 & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \ldots & -2 & 1 & 0 \\ 0 & 0 & 0 & \ldots & 1 & -2 & 1 \end{bmatrix} \in \mathbb{R}^{(n-2)\times n} \tag{2.6}$$

since the filtered series should reflect the 'trend' of the index that is determined by the slope of the piecewise linear function. In fact, Tibshirani (2014) has shown that a second-order trend-filtering analysis on the index is equivalent to a first-order trend-filtering analysis on the return series.

Finally, we will discuss the implementation of the trend-filtering algorithm. The optimisation problem (2.2) is a convex problem since it is sum of norms of affine functions, under the assumption that $\lambda > 0$. The function is also strictly convex (given the existence of the quadratic function), therefore the solution β will be unique, which is known as the trend-filtering estimator.

There are two hyper-parameters in the algorithm. In order to obtain a desired degree of polynomial, we just need to change the value of k. The value of λ can be determined by machine learning techniques such as cross-validation. In practice, we will need to balance between precision and the number of regime switches. Since our asset allocation decision is based on the regime setting, we favour fewer switches in regimes to reduce transaction and market impact costs.

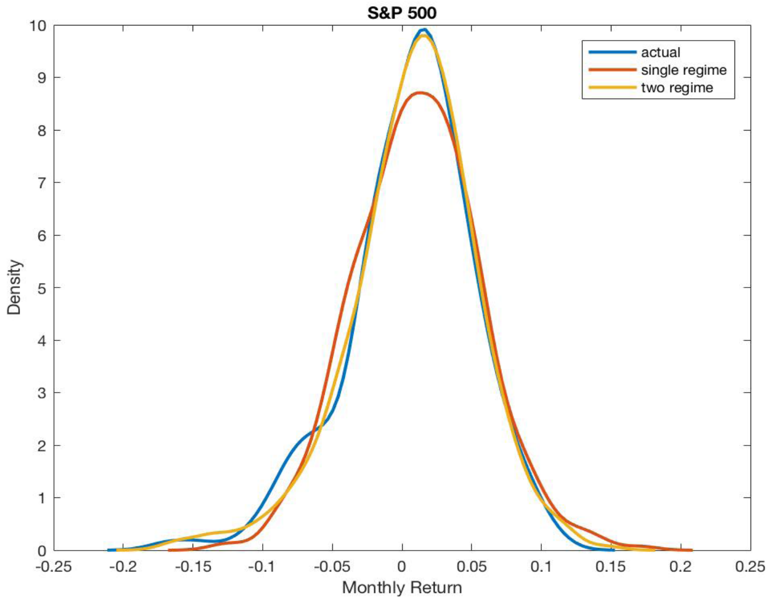
3 Empirical tests

This section describes the performance of asset categories under differing regimes. We apply the trend-filtering algorithm to two sets of historical wealth paths. First, we employ the S&P 500 total return index to separate periods of growth (normal regime) from periods with decreasing equity returns (contraction or crash). Since 1998, about 80% of the months are labelled as normal and 20% contraction.

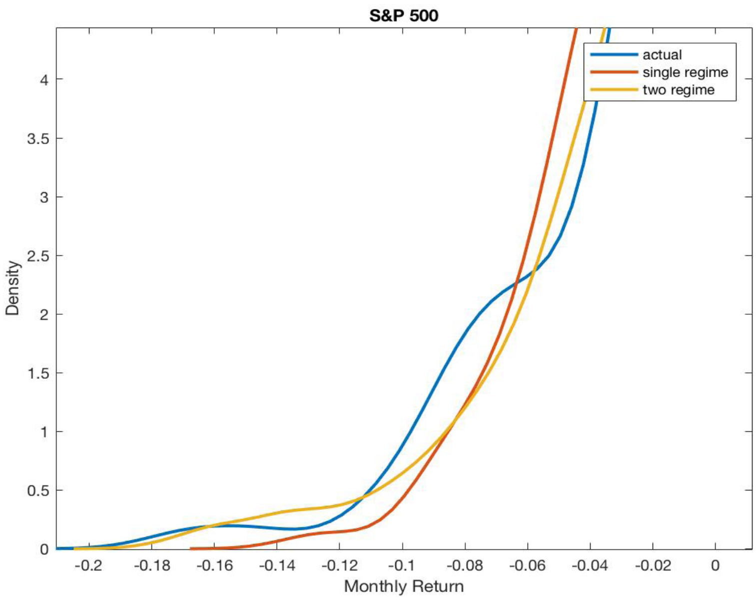
An important advantage of a multi-regime approach to long-term financial planning involves the modelling of the worst-case events. The standard approach is to apply a multi-normal distribution to generate scenarios for use by financial planning systems, including many of the recent robo-advisor platforms. However, this method can underestimate the downside risks. Figures 1(a) and 1(b) show a simple example: a probability density function for the total return of the S&P 500 index monthly over the period 1998 to 2016 and two constructed probability density functions – one for the single regime case and one for a two-regime case. In particular, the left tails are wider for the two-regime model and closer to the actual return distribution than with a single-regime approach. In this example, we apply the generalised lambda distribution for modelling the contraction regime. See Chalabi et al. (2010), Corlu and Corlu (2015) and Karian and Dudewicz (2000) for further details regarding the generalised lambda distribution for addressing the kurtosis and negative skewness of financial data.

Quantile-quantile (QQ) plots in Figures 2(a) and 2(b) display a similar conclusion. In these graphs, it can be seen that the empirical data is left skewed, it has a fat left tail. Looking at the empirical data with single [Figure 2(a)] and two-regimes [Figure 2(b)] simulations, it is obvious that two-regimes simulation provides a better fit than single regime analysis, especially for fat left tails.

Figure 1 (a) Probability density function for S&P 500 – 1998 to 2016* (b) Left tail of probability density function for S&P 500 – 1998 to 2016 (see online version for colours)



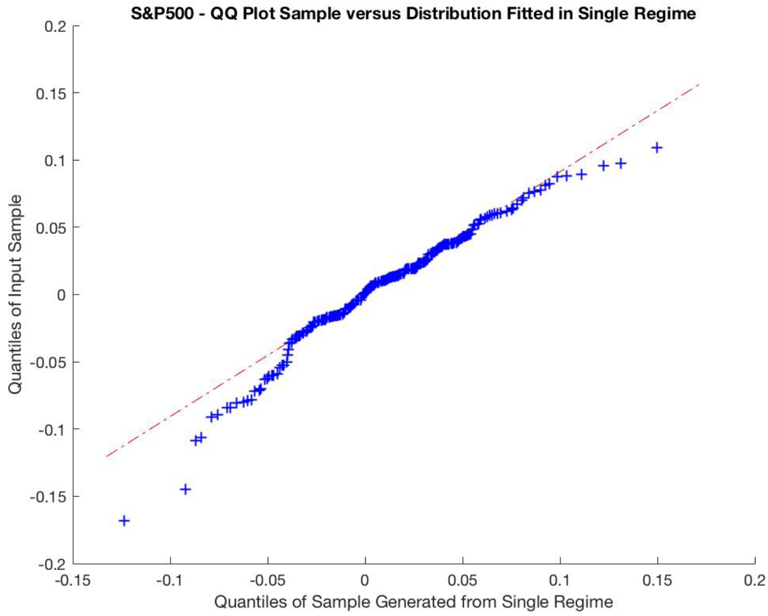
(a)



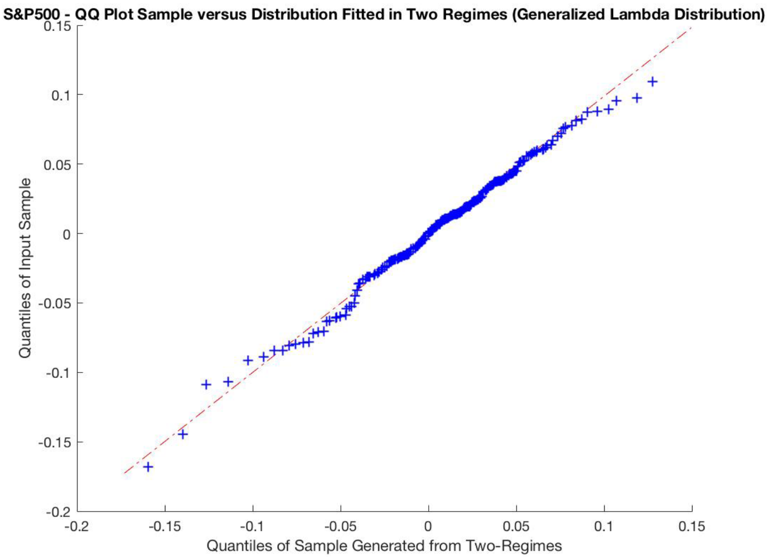
(b)

Note: *Actual returns, simulated returns with single regime and two regimes.

Figure 2 (a) QQ plot for realised S&P 500 versus, (a) single regime simulation (b) double regime simulation (see online version for colours)



(a)



(b)

As we will see in the next section, it is straightforward to include multiple regimes in a scenario-based portfolio model.

In the second analysis, we form regimes based on the notion of inflationary expectation. The requisite time series is the difference between the return of treasury inflation protected bonds (TIPS) minus the return of nominal government bonds. Regimes are thereby {increasing inflationary expectations; and decreasing inflationary expectations}. In both cases, we observe that the performance is highly dependent upon regime and insights regarding diversification follow.

3.1 Regimes via S&P 500

Herein, we analyse the performance of eight core asset categories (US equity, international equity, US treasury, corporate bond, real estate, commodity, TIPS and risk-free) from January 1973 to December 2014 under regimes separated based on S&P 500 index. The category US inflation linked bonds (TIPS) was introduced in 1997; we employ the methodology introduced in Kothari and Shanken (2004) and Martellini et al. (2016) to backfill TIPS whenever its data are unavailable. In addition, as factor investing has been a popular topic in both academia and industry, we will also present factor exposures based on a shrinkage and cross validation methods as described in Blyth et al. (2016). Due to space limitations, we only present empirical results for asset categories in the main text. Results for equity micro-factors and fixed income micro-factors are shown in Appendix. All performance numbers are adjusted for inflation.

To start, we list performance under a single regime as reference. Figure 3 displays the annualised geometric mean of real returns, 5% monthly conditional value-at-risk (CVaR, tail risk) and Sharpe ratio of the eight main asset categories.

From Figure 3, we observe that US and international equity and real estate have the top three geometric mean of real returns, whereas risk-free assets, TIPS and commodity yield the lowest returns. Conversely, real estate, commodity and international equity have the highest monthly tail risks across time and as expected, whereas fixed income securities, such as risk-free, corporate bonds, US treasury and TIPS have lower tail risks. The Sharpe ratio of corporate bonds is higher than for US treasuries, due to the higher volatility of US treasuries. However, as we will see, treasuries provide much better protection during crash periods than corporate bonds due to their flight to quality characteristics.

Next, we consider the underlying factors that drive asset performance. In particular, we conduct a factor-based analysis in conjunction with a robust version that is popular in machine learning. The idea is to apply a shrinkage procedure to reduce the size of the factor loadings in a systematic fashion. There is much evidence that shrinkage approaches are more accurate with out-of-sample tests than via traditional regression methods (Hastie et al., 2009). As with the trend-filtering algorithm, we add a penalty term (regularisation) with a parameter that balances the degree of fit with the size of the beta factor loadings. The size of the penalty parameter is determined by means of a systematic out-of-sample evaluation, called cross validation.

Table 1 displays factor exposures for eight main asset categories and five underlying traded factors using the shrinkage/cross validation (SCV) methodology described in Blyth et al. (2016). The resulting factor loadings are reasonable. First, note the lack of a sizeable number of loadings close to zero. Also, the relationships are consistent with asset pricing models: US equity and international equity have heavy exposures on world equity and small exposures on other factors, for example, US stocks are mainly US companies traded in US dollars, so it has a mild positive exposure on currency protection, while international equity has a negative one. US treasury, corporate bond and TIPS are fixed income, so they have quite heavy exposure on US treasury. TIPS, as expected, has a loading of 0.96 on inflation protection, because it is designed to protect investors against inflation. Commodities have a strong positive loading on inflation and a negative loading on the value of the dollar. In subsequent study, we observe that the driving factors are different under the designated regimes, which helps explain the differential performance of the assets.

Figure 3 Risk and return profiles of asset categories (1973–2014) – inflation adjusted (see online version for colours)

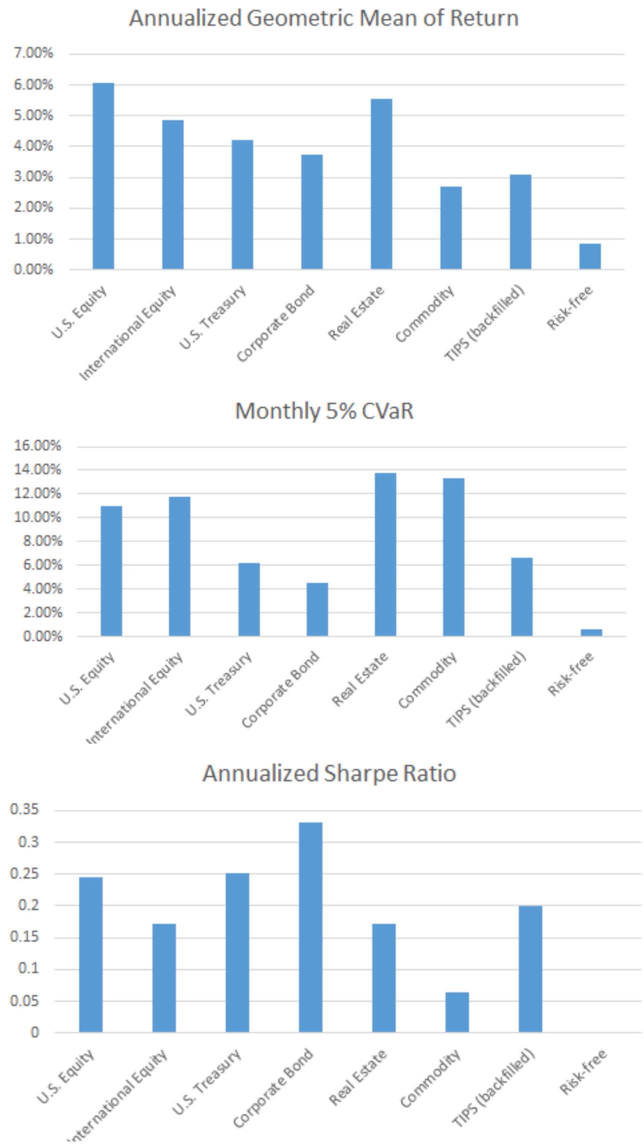


Table 1 Factor exposures for asset categories (January 1973–December 2014)

	(Intercept)	*World equity*	*US treasury*	*High yield*	*Inflation protection*	*Currency protection*
US equity	0.00	0.97	0.15	0.06	0.13	0.50
Intl. equity	0.00	0.96	–0.16	0.00	–0.08	–0.41
US treasury	0.00	0.00	1.35	0.02	0.00	0.10
Corporate bond	0.00	0.07	0.58	0.09	0.08	–0.02
Real estate	0.00	0.69	0.35	0.07	0.30	0.00
Commodity	0.01	0.21	0.00	–0.16	0.59	–0.77
TIPS	0.00	0.00	0.96	0.00	0.96	0.00
Risk free	0.00	0.00	0.00	0.00	0.00	0.00

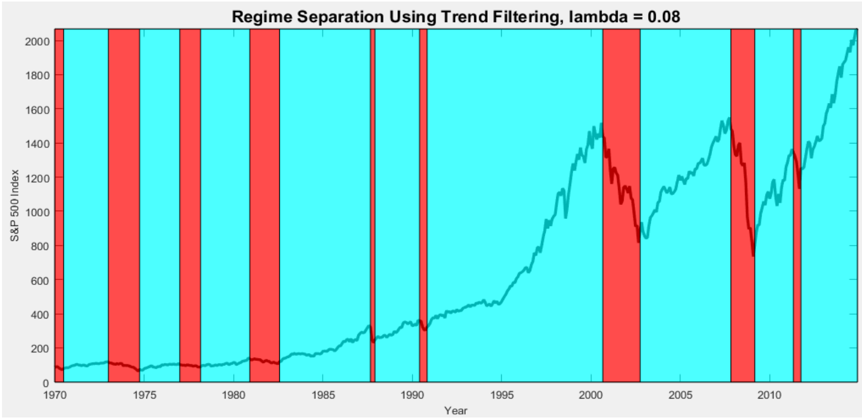
In a traditional portfolio model, we select an historical time period for assisting with the task of setting the parameters in the portfolio model. Generally, these models assume a static economic environment. In contrast, inspired by Ilmanen et al. (2014), who analyse returns of assets under differing macroeconomic conditions, we introduce two approaches to identify economic regimes over historical time periods and analyse how assets perform differently under these regimes. There are, of course, many ways to separate regimes. The examples here serve to illustrate the advantages of regime-based analysis and subsequent asset allocation. As we will see, the variations of risk and return profile and dynamic correlation structure among assets under different regimes are sources of insights for an investor building a portfolio model.

3.2 Regimes separation based on S&P 500

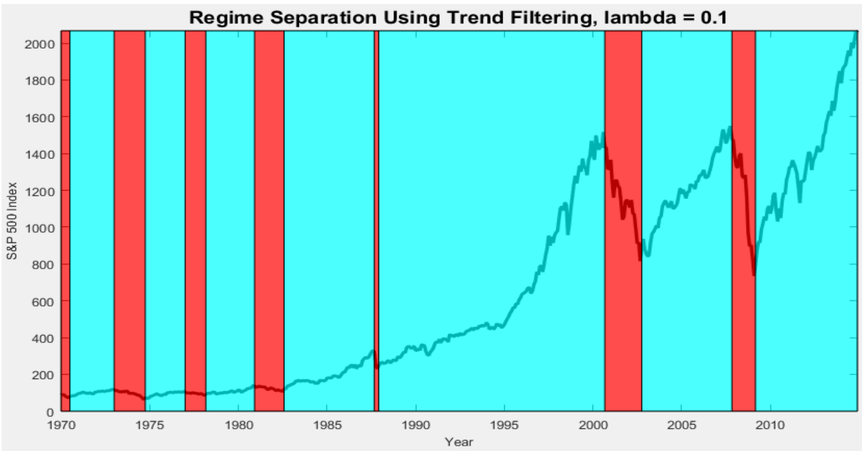
Using monthly adjusted closing prices of S&P 500 from January 1970 to December 2014, we select the penalty parameter $\lambda = 0.08$ to identify two regimes for S&P 500: normal and crash. The black series shown in Figure 4 is S&P 500 index. Cyan denotes normal periods and red denotes crash periods.

The selection of the penalty parameter is a compromise. If the parameter is too small, it will include excess noise and misidentify unnecessary regime switches, whereas if the parameter it too large it will miss important historical crashes (see the previous section for properties of trend-filtering method). By trial and error, $\lambda = 0.08$ is a reasonable choice because it correctly identifies both long bear market periods such as oil crisis before 1975, the internet bubble in early 2000s and the financial crisis in 2008–2009 and short market crash such as the well-known 1987 crash, which is considered to be triggered by various causes such as program trading and market illiquidity. A higher value of λ, such as $\lambda = 0.10$, misses the stock market crash during early 1990s due to oil price increases [see Figure 4(b)]. A lower value of λ, however, makes the model very sensitive to changes of S&P 500: Figure 4(c) shows the regime identification of S&P 500 with $\lambda = 0.06$. Compared to the case where $\lambda = 0.08$, it identifies two more short crash periods: late 1975 and mid-2010. The late-1975 has no record of recession and while the May 2010 flash crash of US stock market was recorded, the event was more like an investor behaviour rather than an economic driven crash, so we believe that it is unreasonable to include it as an economic regime. Therefore, we believe that $\lambda = 0.08$ is a reasonable choice to identify regimes for S&P 500.

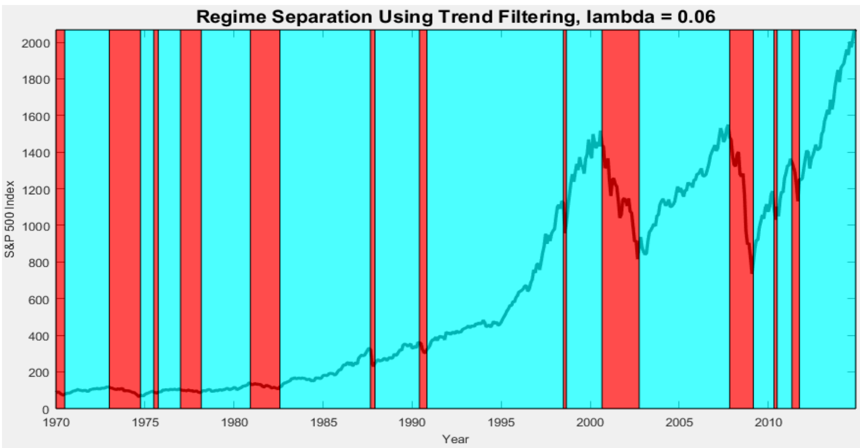
Figure 4 Regime separation of S&P 500 index with (a) $\lambda = 0.08$ (b) $\lambda = 0.10$ and (c) $\lambda = 0.06$ (see online version for colours)



(a)

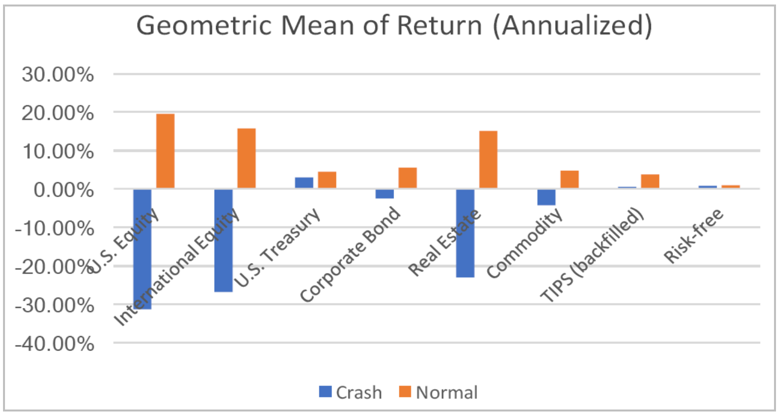


(b)

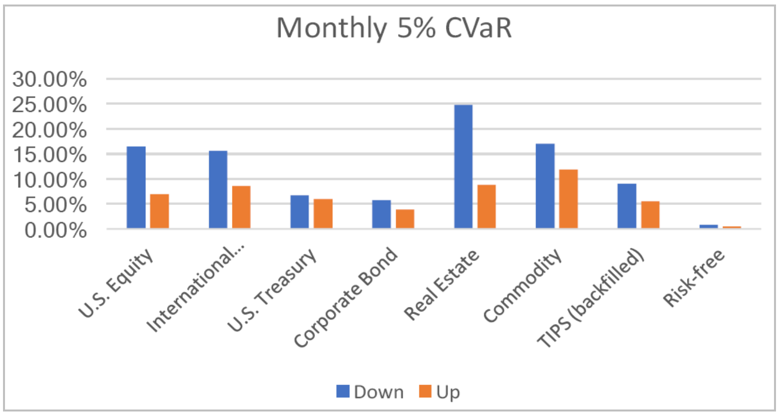


(c)

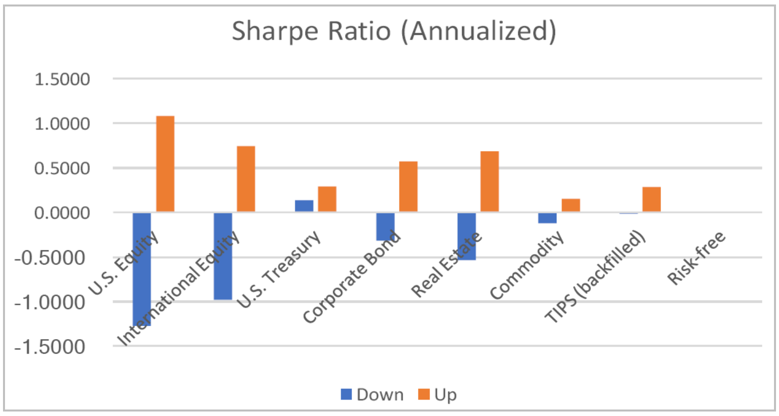
Figure 5 Annualised geometric mean of real return for asset categories (b) Monthly 5% CVaR (c) Annualised Sharpe ratios for asset categories (see online version for colours)



(a)



(b)



(c)

Note: Two regimes, separated by S&P 500.

Next, based on regimes separated above, we analyse performances of asset categories. Figure 5(a) shows the geometric mean of real returns of all asset categories in two regimes. As expected, all asset categories perform much better when S&P 500 index goes up than when it goes down. Equities and real estate yield annualised real returns of over 10% when the market is in normal regime. We also observe that US treasury, TIPS and risk-free assets yield positive real returns even the stock market goes down, suggesting that investors should put capital in fixed income securities as havens during crash periods.

Figure 5(b) lists annualised 5% CVaR in a regime setting. When the stock market goes down, real estate, commodity, equities have the highest tail risks. Investing in fixed income assets is much safer. When the market is in normal period, the tail risks are lower for all asset categories, as expected.

Figure 5(c) shows the annualised Sharpe ratios of all asset categories. Equity assets, real estate and corporate bonds have higher Sharpe ratios than other assets when the market goes up. Interestingly, only US treasury has positive Sharpe ratio when the stock market contracts.

In general, investors should invest in high quality fixed income assets when the market crashes. Since usually the market crashes abruptly, we suggest investors to monitor predictors of market crash and whenever these indicators send alarm signals, they should tilt their portfolio toward fixed income. This strategy is dubbed duration enhancing overlay (DEO) and it is especially pertinent for define-benefit pension plans as discussed in Mulvey et al. (2010).

Table 2 Factor exposures for asset categories

Crash	*(Intercept)*	*World equity*	*US treasury*	*High yield*	*Inflation protection*	*Currency protection*
US equity	0.00	0.97	0.09	–0.05	0.29	0.42
Intl. equity	0.00	0.96	–0.17	0.05	–0.28	–0.48
US treasury	0.00	0.06	1.33	0.01	–0.29	0.04
Corporate bond	0.00	0.05	0.65	0.13	0.42	–0.02
Real estate	0.01	0.82	0.00	0.00	0.00	–0.43
Commodity	–0.01	0.00	0.00	0.00	1.51	–1.21
TIPS	0.00	0.00	0.96	0.00	0.95	0.00
Risk free	0.00	0.00	0.00	0.00	0.01	–0.01
Normal	*(Intercept)*	*World equity*	*US treasury*	*High yield*	*Inflation protection*	*Currency protection*
US equity	0.00	0.91	0.16	0.17	0.07	0.46
Intl. equity	0.00	0.99	–0.18	–0.03	–0.04	–0.38
US treasury	0.00	–0.02	1.35	0.01	0.04	0.11
Corporate bond	0.00	0.05	0.55	0.04	0.00	–0.01
Real estate	0.00	0.53	0.31	0.40	0.22	0.16
Commodity	0.01	0.24	–0.07	–0.39	0.27	–0.57
TIPS	0.00	0.00	0.97	0.00	0.97	0.00
Risk free	0.00	0.00	0.00	0.00	0.00	0.00

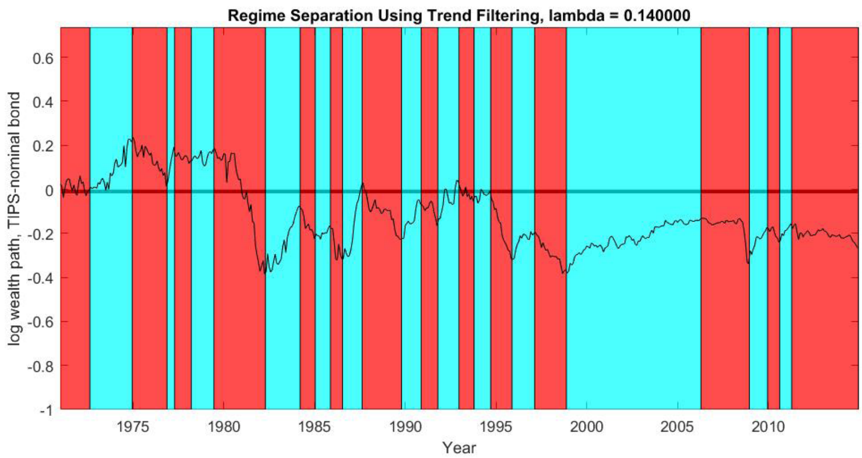
Note: Two regimes, separated by S&P 500.

Table 2 shows factor exposures for asset categories under the two regimes separated by the S&P 500 index. We notice that a few asset categories have much different factor loadings under the two regimes. For example, under the normal environment, real estate has positive exposures on US treasury and high yield bonds, but it has zero exposure on these factors in crash regime. Commodities have positive loadings on world equity and inflation protection during normal regimes, but much lower during contraction periods. Asset such as TIPS, on the contrary, has consistent exposures on US treasury and inflation protection in both regimes. These differential loadings can be included in factor-based asset allocation and ALM studies and they help explain the differential performance of assets under the regimes.

3.3 Regimes identified via expected inflation

In this section, we analyse asset performance based on regimes separated by expected inflation. Expected inflation is calculated by the differences between TIPS return and nominal government bond returns on a monthly basis. For the trend-filtering algorithm, we take penalty parameter $\lambda = 0.14$ over the time period from January 1971 to December 2014 (Figure 6).

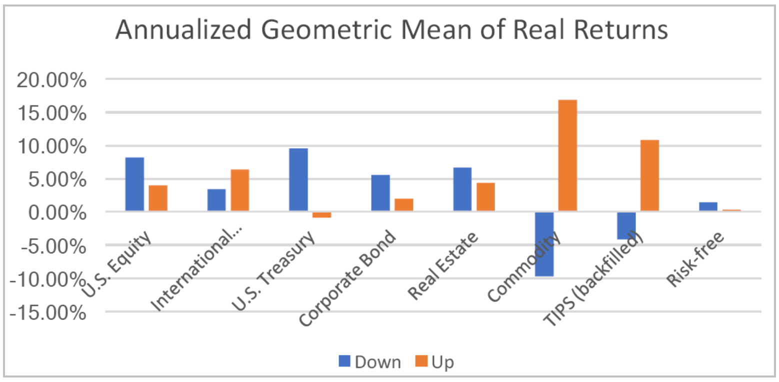
Figure 6 Regime separation of inflationary expectations – 1971 to 2014 (see online version for colours)



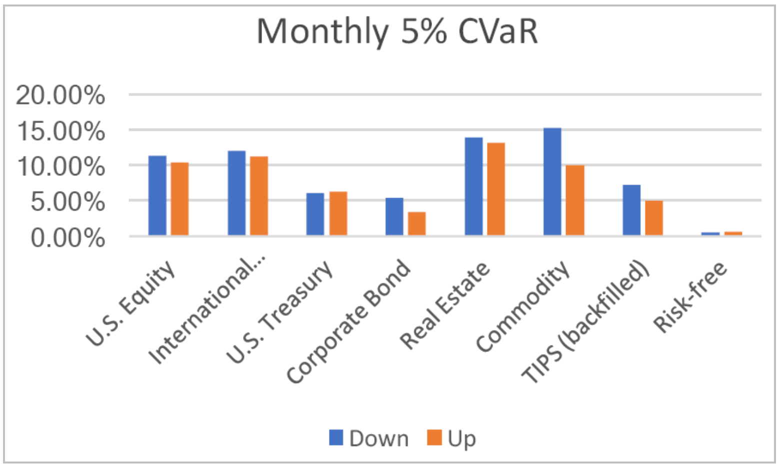
Note: Parameter $\lambda = 0.14$.

Figure 7(a) provides the annualised geometric mean of return of asset categories under two regimes. When expected inflation goes up, all assets, except US treasury, yield positive returns, among which returns of commodity and TIPS are the highest. When expected inflation drops, commodity and TIPS yield negative returns, while US equity, US treasury and real estate give the highest returns. Commodity and TIPs are especially prone to large differences under inflationary expectations, much more than is indicated by a single regime analysis.

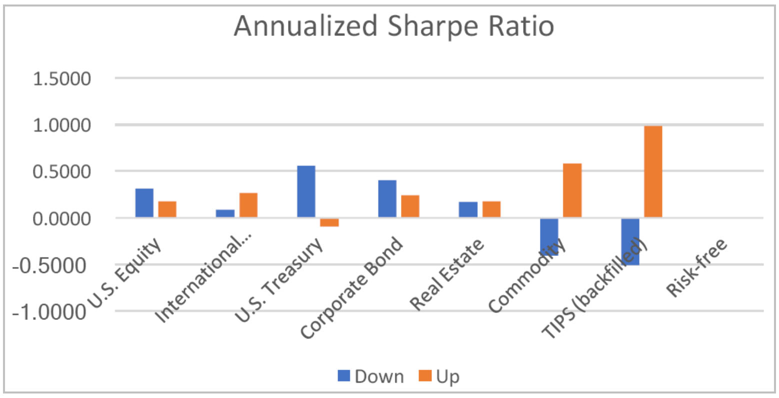
Figure 7 Annualised geometric mean of real return for asset categories (b) Monthly 5% CVaR (c) Annualised Sharpe ratios (see online version for colours)



(a)



(b)



(c)

Note: Two regimes, separated by expected inflation.

Figure 7(b) presents the monthly 5% CVaR of all asset categories. Assets other than fixed income have annualised 5% CVaR around 10% in absolute values when expected inflation goes down. The tail risks of equity assets, US treasury, real estate and risk-free asset do not differ much between two regimes, but commodities and TIPS have significant lower tail risks when expected inflation goes up.

Now, let's look at the Sharpe ratios. Based on Figure 7(c), commodity and TIPS have the highest Sharpe ratio when expected inflation goes up, but they yield negative Sharpe ratio when expected inflation goes down. In contrast, all other assets give positive Sharpe ratios when expected inflation drops. Comparing with Figure 3, in which commodity is not desirable in any performance metric, we find that commodity have excellent returns when expected inflation increases. This illustrates that analysing asset performances under different economic regimes may reveal hidden pattern that might not be discovered in a single regime setting.

Table 3 shows factor exposures for asset categories under regimes separated by expected inflation. Like Table 2, some assets, such as TIPS, do not change factor exposures across regimes, while other assets do. Take commodity as an example, when expected inflation goes up, it has negative exposure on high yield and a 0.25 exposure on inflation protection; however, it has zero exposure on high yield and 0.62 exposure on inflation protection when expected inflation drops.

Table 3 Factor exposures for asset categories

Down	*(Intercept)*	*World equity*	*US treasury*	*High yield*	*Inflation protection*	*Currency protection*
US equity	0.00	1.00	0.10	0.02	0.13	0.51
Intl. equity	0.00	0.93	–0.18	0.09	–0.08	–0.36
US treasury	0.00	–0.04	1.38	0.04	0.00	0.11
Corporate bond	0.00	0.09	0.62	0.12	0.09	0.00
Real estate	0.00	0.89	0.76	0.00	0.63	0.10
Commodity	0.00	0.20	0.00	0.00	0.62	–0.74
TIPS	0.00	0.00	0.97	0.00	0.96	0.00
Risk free	0.00	0.00	0.00	0.00	0.00	0.00
Up	*(Intercept)*	*World equity*	*US treasury*	*High yield*	*Inflation protection*	*Currency protection*
US equity	0.00	0.91	0.20	0.13	0.15	0.48
Intl. equity	0.00	0.98	–0.08	–0.03	0.00	–0.43
US treasury	0.00	0.04	1.31	0.00	0.00	0.07
Corporate bond	0.00	0.04	0.49	0.03	0.00	–0.04
Real estate	0.01	0.44	0.04	0.17	0.00	–0.07
Commodity	0.01	0.19	0.00	–0.32	0.25	–0.58
TIPS	0.00	0.00	0.97	0.00	0.97	0.00
Risk free	0.00	0.00	0.01	–0.02	0.01	0.01

Note: Two regimes, separated by expected inflation.

In summary, in this part, we find that assets and factors can possess differential risk and return profiles under economic regimes and such regimes can be identified by reasonable economic indicators in conjunction with machine learning approaches. Since market structure keeps evolving over time and the dynamics of risks and returns are changing, our analysis based on static historical data cannot be the sole determinate for future investments. Still, investing based on economic regimes and factors can help us better understand the dynamics of asset performances.

Further exploration is needed in discovering risk and return profiles for asset categories and micro-factors under regimes separated by different macroeconomic conditions. We believe that the portfolios constructed based on regime-based performances of assets and micro-factors are at least as good as the ones without regimes. In next part, we will construct portfolios based on mean-variance optimisation and mean-CVaR optimisation to illustrate this idea.

4 Scenario-based asset allocation and ALM models

In this section, a regime-based asset allocation model is built and compared with the traditional mean-variance portfolio. The development of a scenario-based portfolio model is quite general, allowing for almost any type of distribution or range of events. The simplest approach is to assume a multi-normal distribution $N(\mu, Q)$ for mean returns and a covariance matrix and to sample this distribution to generate scenarios. The basic scenario-based portfolio model has the following structure:

1 Generate $m = 10{,}000$ scenarios as a set $\{S\}$ based on $N(\overline{r}, Q)$ with historical mean $\overline{r}$ and covariance matrix Q.

2 Solve the following Markowitz portfolio optimisation problem for target return r_{target}:

$$\begin{aligned} &\text{minimise } \sigma_s = \text{stdev}(r_s) \\ &\qquad \text{s.t.} \quad 0 < \text{w}_i < 1 \;\; (w_i \text{ is the weight for asset } i) \\ &\qquad\qquad \sum w_i = 1 \\ &\qquad\qquad r_s = \{\boldsymbol{w} \cdot \boldsymbol{s} \mid s \in \{S\}\} \\ &\qquad\quad \text{mean}(r_s) >= r_{target}. \end{aligned}$$

The generator of this model (the first step) is easily adaptable for multiple regimes:

1'

- Run the regime analysis and compute the percentage of time in normal regime (d_{normal}), the historical mean and covariance for two regimes $(\overline{r}_{normal}, \overline{r}_{crash}, Q_{normal}, Q_{crash})$.
- Generate md scenarios as a set $\{S_1\}$ based on normal distribution with historical mean $\overline{r}_{normal}$ and covariance matrix Q_{normal}.
- Generate $m(1 - d)$ scenarios as a set $\{S_2\}$ based on normal distribution with historical mean $\overline{r}_{crash}$ and covariance matrix Q_{crash}.

As discussed, the traditional mean-variance portfolio optimisation can be readily modified to address the worst-case scenarios generated by the two-regime analysis above.

One possible model involves minimising the CVaR. CVaR, as a consistent measure of risk, is developed upon value-at-risk (VaR):

$$\mathrm{VaR}_h(X) = -\min\{c \mid P(X \le c) \ge h\}.$$

This VaR at a specific probability level h is defined as a number v such that with only probability h, the amount of loss in asset X will exceed this v. VaR reflects only the point where the worst case with probability h happens, so we introduce CVaR to reflect the average loss of these worst cases:

$$\begin{aligned}\mathrm{CVaR}_h(X) &= -E\left[X \mid X \le \mathrm{VaR}_h(X)\right]\\ &= \mathrm{VaR}_h(X) + \frac{\int_{\mathbb{R}}\left[-X - \mathrm{VaR}_h(X)\right] + f(x)dx}{1-h}.\end{aligned}$$

Rockafellar and Uryasev (2000) provide further details of optimising under VaR and CVaR via a scenario-based portfolio model. Now, with CVaR as our measure of risk, we can modify the portfolio optimisation problem as:

2’ Solve the following mean-CVaR portfolio optimisation problem for target return r_{target}:

$$\begin{aligned}&\text{minimise } \mathrm{CVaR}_h\\ &\quad \text{s.t.} \quad 0 < \mathrm{w}_i < 1 \ \left(w_i \text{ is the weight for asset } i\right)\\ &\qquad \sum w_i = 1\\ &\qquad r = \left\{\boldsymbol{w}\cdot\boldsymbol{s} \mid s \in \{S_1\} \cup \{S_2\}\right\}\\ &\qquad E[r] >= r_{target}.\end{aligned}$$

If we construct auxiliary variables v for VaR and u_j for excess loss under scenario j ($j = 1, \ldots, m$), we can rewrite the problem as a linear program:

$$\begin{aligned}&\text{minimise } v + \frac{\sum_{j=1}^{m} u_j}{m(1-h)}\\ &\text{subject to } \sum w_i = 1,\ w_i \ge 0\\ &\qquad E[r] \ge r_{target}\\ &\qquad u_i \ge 0,\ r^i + v + u_i \ge 0.\end{aligned}$$

Combine the new Steps 1’ and 2’, we obtain a portfolio model that is scenario-based and more adaptable to the minimisation of tail risks.

Based on the methods introduced above, we construct portfolios with empirical data. Parameters of portfolios are estimated from historical data from 1973 to 2006 and portfolios are evaluated based on data from 2007 to 2015 in an out-of-sample analysis. The data include eight asset categories (US equity, international equity, US treasury, corporate bond, real estate, commodity, TIPS and risk-free) and the data sources are listed in Appendix. All returns are in real terms, accounted for inflation as computed from CPI index.

According to the return and risk analysis in Figure 3, we set target real returns from 3.5% to 6% annually, with an increment of 0.5%. Only private equity achieved an average real return of more than 5.5%, so the 6% target will be mostly invested in private equity in all portfolios. The lower bound, 3.5%, is set as the average of bond returns.

The goal of these out of sample empirical tests is to evaluate the performance of portfolios during the 2007–2009 global crash (with maximum drawdown as criteria) and the subsequent performance until 2015. As such, we do not rebalance or re-optimise during the evaluation period. The consideration is that rebalancing may hurt the portfolio during 'crash' periods when the asset prices witness persistent drawdown.

Tables 4, 5 and 6 shows standard risk and return evaluation metrics for mean-variance portfolios (one regime), mean-variance portfolios (two regimes) and mean-CVaR portfolios (two regimes).

Table 4 Portfolio evaluation

Portfolio target return	*3.50%*	*4.00%*	*4.50%*	*5.00%*	*5.50%*	*6.00%*
Geo-mean	0.08%	0.12%	0.15%	0.16%	0.11%	0.08%
Geo-mean (annualised)	1.00%	1.43%	1.80%	1.98%	1.36%	0.95%
VaR (5%)	4.34%	5.28%	6.49%	7.78%	10.51%	13.37%
CVaR (5%)	5.73%	7.32%	9.06%	11.25%	15.83%	21.44%
Volatility (monthly)	2.10%	2.65%	3.23%	3.93%	5.29%	6.85%
Volatility (annualised)	7.27%	9.18%	11.19%	13.63%	18.33%	23.74%
Sharpe ratio (monthly)	0.0122	0.0131	0.0134	0.0126	0.0079	0.0049
Sharpe ratio (annualised)	0.0421	0.0454	0.0465	0.0437	0.0274	0.0171
Max. drawdown	24.46%	29.83%	34.98%	40.81%	50.39%	58.38%

Note: Mean-variance portfolios.

Table 5 Portfolio evaluation

Portfolio target return	3.50%	4.00%	4.50%	5.00%	5.50%	6.00%
Geo-mean	0.09%	0.12%	0.16%	0.18%	0.11%	0.04%
Geo-mean (annualised)	1.06%	1.51%	1.92%	2.19%	1.37%	0.52%
VaR (5%)	3.66%	4.42%	5.18%	6.18%	9.85%	13.61%
CVaR (5%)	5.03%	6.40%	7.83%	9.49%	14.68%	21.97%
Volatility (monthly)	1.82%	2.30%	2.79%	3.37%	4.97%	7.00%
Volatility (annualised)	6.32%	7.95%	9.65%	11.66%	17.23%	24.24%
Sharpe ratio (monthly)	0.0139	0.0148	0.0151	0.0146	0.0083	0.0028
Sharpe ratio (annualised)	0.0480	0.0512	0.0523	0.0506	0.0287	0.0096
Max. drawdown	21.39%	26.12%	30.67%	35.82%	48.31%	59.04%

Note: Mean-variance portfolios, two regimes.

Table 4 shows that geometric return for portfolios increase when target returns increase from 3.5% to 5% and then decrease afterwards. Risk metrics (VaR, CVaR, volatility and maximum drawdowns) increase when target returns increase because when target return is higher, investment in assets that provide high returns and high risks, such as equity, is higher and downside risks inevitably increase. Sharpe ratio for portfolio with target return of 4.5% is the highest.

Table 6 Portfolio evaluation

Portfolio target return	3.50%	4.00%	4.50%	5.00%	5.50%	6.00%
Geo-mean	0.10%	0.14%	0.18%	0.19%	0.13%	0.01%
Geo-mean (annualised)	1.17%	1.68%	2.13%	2.35%	1.54%	0.08%
VaR (5%)	3.42%	4.06%	4.74%	6.18%	9.40%	13.92%
CVaR (5%)	4.88%	6.17%	7.53%	9.44%	14.86%	23.30%
Volatility (monthly)	1.76%	2.20%	2.67%	3.32%	4.96%	7.31%
Volatility (annualised)	6.09%	7.62%	9.25%	11.50%	17.19%	25.33%
Sharpe ratio (monthly)	0.0150	0.0160	0.0163	0.0154	0.0093	0.0004
Sharpe ratio (annualised)	0.0520	0.0554	0.0563	0.0535	0.0322	0.0014
Max. drawdown	20.46%	24.82%	29.15%	35.14%	47.88%	60.32%

Note: Mean-CVaR portfolios, two regimes.

Comparing to Table 4, portfolios in Table 5 have better risk and return performances except for the portfolio with target return of 6%. Take portfolios with 4% target return as an example, portfolio in Table 5 yields higher geometric mean of return (1.51% vs. 1.43%, annualised), lower 5% VaR (4.42% vs. 5.28), higher Sharpe ratio (0.0512 vs. 0.0454, annualised) and lower maximum drawdown (26.12% vs. 29.83%). Such results are as expected because portfolios formed in Table 5 are constructed based on returns with mixed normal distributions, which include information of two different economic regimes. Comparatively speaking, regime-based portfolio construction captures information of different regimes and thus can adapt to different economic environments more easily.

Portfolios shown in Table 6 are also constructed in a regime-based setting. Different from those in portfolios in Table 5, these portfolios are optimised in a mean-CVaR context, which focuses on protecting against downside risks. As we can see, the mean-CVaR portfolios with target returns from 3.5% to 5.0% outperform corresponding mean-variance portfolios in all metrics (higher returns, lower tail risks, lower volatility, higher Sharpe ratio and lower maximum drawdown). The mean-CVaR portfolio with target return 5.5% also beat the mean-variance portfolio in Table 5 in almost all metrics except for 5% CVaR (14.86% for mean-CVaR portfolio vs. 14.68% for mean-variance portfolio). The portfolio with target return of 6% has inferior performance than the mean-variance portfolio.

In summary, in our example, we observe that traditional mean-variance portfolio optimisation methods and related downside risk measurements tend to underestimate the frequency of worst-case events. Accordingly, we can improve the downside risk estimation through a regime-based approach. Furthermore, for the example, mean-CVaR optimisation with multiple regimes is superior to mean-variance optimisation in the way that it leads to portfolios that protect investors against severe loss during sharp drawdown periods.

Importantly, the analysis of ‘worst-case’ events is limited by the historical evidence. There is always the chance of a massive political crisis such as a shooting war, leading to a severe meltdown in economic markets. Other dire events might be triggered by a trade war between the USA and China, or another terrorist attach similar in consequence to September 11, 2001. The abiding concept in financial economics is that investors are paid to take on risks and greater overall risks lead to higher average rewards. Nevertheless, there is much evidence that many investors are susceptible to neglected risks and these risks are exposed during market crashes. It is wise, therefore, to consider tail risks when constructing portfolios for investors with long time horizons. The topics discussed in this paper can be extended in a natural way for asset and liability management systems.

5 Conclusions

In this paper, we introduced the trend-filtering algorithm for regime-separation and applied it to examine the risk and return profile of assets and factors under different regimes. Based on the regime separation of the stock market, we conducted a two-regime mean-CVaR portfolio optimisation to exhibit the benefit of the portfolio construction techniques that consider crash events.

Importantly, the analysis of 'worst-case' events is limited by the historical evidence. There is always the chance of a massive political crisis such as a shooting war, leading to a severe meltdown in economic markets. Other dire events might be triggered by a trade war between the USA and China, or another terrorist attach similar in consequence to September 11, 2001.

The abiding concept in financial economics is that investors are paid to take on risks and greater overall risks lead to higher average rewards. Nevertheless, there is much evidence that many investors are susceptible to neglected risks and these risks are exposed during market crashes. It is wise, therefore, to consider tail risks when constructing portfolios for investors with long time horizons. The topics discussed in this paper can be extended in a natural way for asset and liability management systems.

Appendices/Supplementary materials are available on request by emailing the corresponding author.

References

Ang, A. and Bekaert, G. (2002) 'International asset allocation with regime shifts', *The Review of Financial Studies*, Vol. 15, No. 4 pp.1137–1187.

Bae, G., Kim, W.C. and Mulvey, J. (2013) 'Dynamic asset allocation for varied financial markets under regime switching framework', *European Journal of Operational Research*, Vol. 234, No. 2, pp.450–458.

Blyth, S., Szigety, M.C. and Xia, J. (2016) 'Flexible indeterminate factor-based asset allocation', *Journal of Portfolio Management*, Special Issue, Vol. 42, No. 5, pp.79–93.

Chalabi, Y., Scott, D. and Wurtz, D. (2010) *The Generalized Lambda Distribution as an Alternative to Model Financial Returns*, Working Paper, ETH and University of Auckland.

Corlu, C. and Corlu, A. (2015) 'Modelling exchange rate returns: which flexible distribution to use?', *Quantitative Finance*, Vol. 15, No. 11, pp.1851–1864.

Gennaioli, N., Shleifer, A. and Vishny, R. (2012) 'Neglected risk, financial innovation, and financial fragility', *Journal of Financial Economics*, Vol. 104, No. 3, pp.452–468.

Graflund, A. and Nilsson, B. (2003) 'Dynamic portfolio selection: the relevance of switching regimes and investment horizon', *European Financial Management*, Vol. 9, No. 2, pp.179–200.

Guidolin, M. and Timmermann, A. (2007) 'Asset allocation under multivariate regime switching', *Journal of Economic Dynamics and Control*, Vol. 31, No. 11, pp.3503–3544.

Guidolin, M. and Timmermann, A. (2008) 'International asset allocation under regime switching, skew, and kurtosis preferences', *The Review of Financial Studies*, Vol. 21, No. 2, pp.889–935.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, New York.

Ilmanen, A., Maloney, T. and Ross, A. (2014) 'Exploring macroeconomic sensitivities: how investments respond to different economic environments', *The Journal of Portfolio Management*, Spring, Vol. 40, No. 3, pp.87–99.

Karian, Z. and Dudewicz, E. (2000) *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*, Chapman and Hall, Chapman and Hall/CRC, Boca Raton, FL.

Kim, S.J., Koh, K., Boyd, S. and Gorinevsky, D. (2009) 'L1 trend filtering', *SIAM Review*, Vol. 51, No. 2, pp.339–360.

Kothari, S.P. and Shanken, J. (2004) 'Asset allocation with inflation-protected bonds', *Financial Analysts Journal*, February, Vol. 60, No. 1, pp.54–70.

Kyle, A. and Xiong, W. (2001) 'Contagion as a wealth effect', *Journal of Finance*, Vol. 56, No. 4, pp.1401–1440.

Martellini, L., Milhau, V. and Tarelli, A. (2016) 'Hedging inflation-linked liabilities without inflation-linked instruments through long/short investments in nominal bonds', *The Journal of Fixed Income*, December, Winter, Vol. 24, No. 3.

Mulvey, J.M. and Liu, H. (2016) 'Identifying economic regimes: reducing downside risks for university endowments and foundations', *Journal of Portfolio Management*, Fall, Vol. 43, No. 1, pp.100–108.

Mulvey, J.M. and Reus, L. (2016) 'Dynamic allocations for currency futures under switching regimes signals', *European Journal of Operational Research*, Vol. 253, No. 1, pp.85–93.

Mulvey, J.M., Kim, W.C. and Ma, Y. (2010) 'Duration enhancing overlay strategies for defined-benefit pension plans', *Journal of Asset Management*, Vol. 11, Nos. 2–3, pp.136–162.

Rockafellar, R.T. and Uryasev, S. (2000) 'Optimization of conditional value-at-risk', *Journal of Risk*, Spring, Vol. 2, No. 3, pp.21–41.

Tibshirani, R.J. (2014) 'Adaptive piecewise polynomial estimation via trend filtering', *The Annals of Statistics*, March, Vol. 42, No. 1, pp.285–323.