A simple procedure to incorporate predictive models in a continuous time asset allocation

Mark Davis & Sébastien Lleo

Routledge
Taylor & Francis Group

BUBBLES AND CRASHES

A simple procedure to incorporate predictive models in a continuous time asset allocation

MARK DAVIS[†] AND SÉBASTIEN LLEO[‡*]

[†]*Imperial College London, UK*
[‡]*NEOMA Business School, France*

Stochastic optimisation has found a fertile ground for applications in finance. One of the greatest challenges remains to incorporate a set of scenarios that accurately model the behaviour of financial markets, and in particular their behaviour during crashes and crises, without sacrificing the tractability of the optimal investment policy. This paper shows how to incorporate return predictions and crash predictions as views into continuous time asset allocation models.

Keywords: Black–Litterman; Kalman filter; Stochastic control; Risk-sensitive control; Asset management; Expert opinion; Equity market crashes; BSEYD; CAPE

JEL Classification: C11; C13; C61; G11

1. Introduction

Stochastic optimisation has advanced at a remarkable pace over the past 40 years. One of the greatest challenge in stochastic optimisation is to incorporate a set of scenarios that accurately models the behaviour of financial markets without sacrificing the tractability of the optimal investment policy.

Stochastic control, the branch of stochastic optimisation solving problems set in continuous time, often has tractable analytical or numerical solutions. However, most stochastic control models are limited by unrealistic assumptions that fail to capture adequately the behaviour of financial markets.

Recent progress in stochastic controls has created new opportunities to build more realistic models. For example, the development of viscosity solutions has made it possible to develop models with jumps and stochastic volatility. In some instances, viscosity solutions provide the key argument to prove that jump-diffusion models admit a smooth solution (Davis *et al.* 2010, Davis and Lleo 2013b).

Another important milestone is the development of stochastic control models inspired by the Black–Litterman model (Black and Litterman 1990, 1991, 1992). These stochastic control models incorporate views formulated by securities and market analysts into the parameter estimation process to produce forward-looking scenarios that a simple historical analysis would not consider (see Frey *et al.* 2012, Davis and Lleo 2013a, Gabih *et al.* 2014a,b, Davis and Lleo 2015, 2016).

Berge *et al.* (2008) showed that the ability to forecast market corrections and crashes improves significantly the risk-adjusted performance of long-term investors. Incorporating views on whether crashes are likely to occur is therefore highly relevant to investment managers. In this paper, we show how to incorporate return prediction and crash prediction models as views in the Black–Litterman model

in continuous time proposed by Davis and Lleo (2013a). We look at two specific examples of return prediction and crash prediction models: the Cyclically Adjusted Price-to-Earnings ratio (CAPE) and the Bond-Stock Earnings Yield Differential (BSEYD) model.

2. Black–Litterman in continuous time

Developed by Davis and Lleo, the Black–Litterman model in continuous time uses linear filtering to incorporate analyst views and expert opinions in a continuous time asset allocation. The key to the approach is that the filtering problem and the stochastic control problem are effectively separable. The continuous time model uses this insight to incorporate analyst views and non-investable assets as observations in the filter even though they are not present in the portfolio optimisation.

The model has four key components: (i) the financial market, (ii) the views, (iii) the linear filter and (iv) the stochastic control problem.

In this paper, we focus only on a market with investable assets in order to keep the discussion clear and concise. The extension to non-investable assets is straightforward, and we refer the reader to Davis and Lleo (2013a) for the details.

2.1. *The financial market: asset prices are driven by unobservable factors*

Start with a financial market comprising $m \geq 1$ risky securities S_i, $i = 1, \ldots m$, and a money market account process S_0. The growth rates of the assets depend on n unobservable factors $X_1(t), \ldots, X_n(t)$ which follow the affine dynamics given in Equation (1).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space. On this space, define an $\mathbb{R}^N$-valued $(\mathcal{F}_t)$-Brownian

---

*Corresponding author. Email: sebastien.lleo@neoma-bs.fr

motion $W(t)$ with components $W_j(t)$, $j = 1, \ldots, N$, and $N := n + m + k$. This is an incomplete market setting with n sources of risks corresponding to the factors, m sources of risks related to the assets and k sources of uncertainty related to the analyst views.

The asset returns and risk premia are subject to the evolution of the n-dimensional vector of unobservable factors $X(t)$ modelled as an affine process

$$\mathrm{d}X(t) = (b + BX(t)) \, \mathrm{d}t + \Lambda \, \mathrm{d}W(t), \quad X(0) = x. \quad (1)$$

The dynamics of the money market asset S_0 is given by

$$\frac{\mathrm{d}S_0(t)}{S_0(t)} = (a_0 + A_0' X(t)) \, \mathrm{d}t, \quad S_0(0) = s_0, \quad (2)$$

and that of the m risky assets follows the Stochastic Differential Equations (SDEs)

$$\frac{\mathrm{d}S_i(t)}{S_i(t)} = (a + AX(t))_i \, \mathrm{d}t + \sum_{j=1}^{N} \sigma_{ik} \, \mathrm{d}W_k(t),$$
$$S_i(0) = s_i, \quad i = 1, \ldots, m. \quad (3)$$

For convenience, we denote by Σ the matrix (σ_{ij}).

The approach assumes that no two assets have the same risk profile:

ASSUMPTION 2.1   *The matrix $\Sigma \Sigma'$ is positive-definite.*

The discounted asset price $\tilde{S}_i(t)$ is

$$\tilde{S}_i(t) = \frac{S_i(t)}{S_0(t)}, \quad i = 1, \ldots, m, \quad \tilde{S}_i(0) = \frac{s_i}{s_0}.$$

The risk premium $\mathfrak{s}_i(t)$ is $\mathfrak{s}_i(t) = \log(\tilde{S}_i(t))$, $i = 1, \ldots, m$. Hence, $\mathfrak{s}_i(t)$ solves the SDE

$$\mathrm{d}\mathfrak{s}_i(t) = \left[(\tilde{a} + \tilde{A}X(t))_i - \frac{1}{2} \Sigma \Sigma'_{ii} \right] \mathrm{d}t + \sum_{k=1}^{N} \sigma_{ik} \, \mathrm{d}W_k(t),$$
$$\mathfrak{s}_i(0) = \log \frac{s_i}{s_0}, \quad i = 1, \ldots, m, \quad (4)$$

where $\tilde{a} = a - a_0 \mathbf{1}$, $\tilde{A} = A - A_0 \mathbf{1}$ and $\mathbf{1} \in \mathbb{R}^m$ denotes the m-element column vector with entries equal to 1. The dynamics of the risk premia is Gaussian (conditional on $X(t)$). We can use risk premia as an observation in a linear filtering.

2.2. Analyst views

Analysts formulate views about risk premia or the spread between risk premia over a time horizon. A typical analyst statement would be:

> my research leads me to believe that the spread between 10-year Treasury Notes and 3-month Treasury Bills will remain low over the next year before gradually widening over the following two years to 200 basis points in response to improving macroeconomic conditions. I am 90% confident that the spread will be within a range of 180 bps to 220 bps in two years.

Mathematically, these statements translate the k views into a system of stochastic differential equations

$$\mathrm{d}Z(t) = (a_Z(t) + A_Z(t)X(t)) \, \mathrm{d}t$$
$$+ \Psi(t) \, \mathrm{d}W(t), \quad Z(0) = z, \quad (5)$$

where $W(t)$ is the N-dimensional Brownian motion and Ψ is a $k \times N$ matrix with zeros on its first $(n + m)$ rows. Calibrate the drift $a_Z(t) + A_Z(t)X(t)$ to the central view, and the diffusion matrix Ψ to the confidence interval around the view. This entire construction takes place at initial time $t = 0$. It involves neither modelling the arrival of new opinions nor the evolution of the views over time.

In the filtering step, we will need to invert the matrix $\Psi \Psi'$. We therefore require the following:

ASSUMPTION 2.2   *The matrix $\Psi \Psi'$ has full rank.*

2.3. Filter the views and asset prices to estimate the factors

We refer the reader to Davis and Lleo (2013a) for the general case $r(t) = a_0 + A_0 X(t)$. Here, we outline the solution when $A_0 = 0$.

There are two sources of observations for the risk premia:

   (i) m investable risky assets $S_1(t), \ldots, S_m(t)$;
   (ii) k analyst views $Z_1(t), \ldots, Z_k(t)$.

The pair of processes $(X(t), Y(t))$, where

$$Y_i(t) = \begin{cases} \mathfrak{s}_i(t) = \log \dfrac{S_i(t)}{S_0(t)}, & i = 1, \ldots, m, \\ Z_{i-M}(t), & i = m + 1, \ldots, m + k \end{cases} \quad (6)$$

takes the form of the 'signal' and 'observation' processes in a Kalman filter system, and consequently the conditional distribution of $X(t)$ is normal $N(\hat{X}(t), P(t))$, where $\hat{X}(t) = \mathbb{E}[X(t) | \mathcal{F}_t^Y]$ satisfies the Kalman filter equation and $P(t)$ is a deterministic matrix-valued function.

Express the dynamics of $Y(t)$ succinctly as

$$\mathrm{d}Y(t) = (a_Y(t) + A_Y(t)X(t)) \, \mathrm{d}t + \Xi(t) \, \mathrm{d}W(t),$$
$$Y(0) = y_0, \quad (7)$$

where the $(m + k)$-element vector a_Y, $(m + k) \times n$ matrix A_Y and the $(m + k) \times N$ matrix Ξ are given by

$$a_Y(t) = \begin{pmatrix} \tilde{a} - \dfrac{1}{2} \Sigma \Sigma \\ a_Z(t) \end{pmatrix}, \quad A_Y(t) = \begin{pmatrix} \tilde{A} \\ A_Z(t) \end{pmatrix},$$

$$\Xi(t) = \begin{pmatrix} \Sigma \\ \Psi(t) \end{pmatrix}.$$

Next, define two processes $Y^1(t), Y^2(t) \in \mathbb{R}^{m+k}$ as follows:

$$Y^1(t) = A_Y(t)X(t) \, \mathrm{d}t + \Psi(t) \, \mathrm{d}W(t), \quad Y^1(0) = 0, \quad (8)$$
$$Y^2(t) = a_Y(t) \cdot \mathrm{d}t, \quad Y^2(0) = y_0, \quad (9)$$

so that $Y(t) = Y^1(t) + Y^2(t)$.

In the present case, assume that X_0 is a normal random vector $N(m_0, P_0)$ with known mean m_0 and covariance P_0, and that X_0 is independent of the Brownian motion W. The processes $(X(t), Y^1(t))$ satisfying Equations (1) and (8) and the filtering equations, which are standard, are stated in the following proposition.

PROPOSITION 2.3 (Kalman filter) *The conditional distribution of $X(t)$ given $\mathcal{F}_t^Y$ is $N(\hat{X}(t), P(t))$, which is calculated as follows.*

(i) *The innovations process $U(t) \in \mathbb{R}^{m+k}$ defined by*

$$dU(t) = (\Xi\Xi')^{-1/2}(dY(t) - A_Y\hat{X}(t)\,dt),$$
$$U(0) = 0 \qquad (10)$$

is a vector Brownian motion.

(ii) *$\hat{X}(t)$ is the unique solution of the SDE*

$$d\hat{X}(t) = (b + B\hat{X}(t))\,dt + \check{\Lambda}(t)\,dU(t),$$
$$\hat{X}(0) = m_0, \qquad (11)$$

where $\check{\Lambda}(t) = (\Lambda\Xi' + P(t)A_Y')(\Xi\Xi')^{-1/2}$.

(iii) *$P(t)$ is the unique non-negative definite symmetric solution of the matrix Riccati equation*

$$\dot{P}(t) = \Lambda\mathfrak{p}^\perp(\mathfrak{p}^\perp)'\Lambda' - P(t)A_Y'(\Xi\Xi')^{-1}A_YP(t)$$
$$+ (B - \Lambda\Xi'(\Xi\Xi')^{-1}A_Y)P(t)$$
$$+ P(t)(B' - A_Y'(\Xi\Xi')^{-1}\Xi\Lambda'), \quad P(0) = P_0,$$

where $\mathfrak{p}^\perp := I - \Xi'(\Xi'\Xi)^{-1}\Xi$.

Now, the Kalman filter has replaced the initial state process $X(t)$ by an estimate $\hat{X}(t)$ with dynamics given in Equation (10). To recover the asset price process, we use Equations (6)–(11) together with Equation (9) to obtain the dynamics of $Y(t)$

$$dY(t) = dY_1(t) + dY_2(t)$$
$$= (a_Y + A_YX(t))\,dt + (\Xi\Xi')^{1/2}\,dU(t),$$
$$Y(0) = y_0, \qquad (12)$$

and from there, we recover the dynamics of $Z(t)$ and $S(t)$.

Observe that $\Xi\Xi' := \left(\begin{smallmatrix}\Sigma\Sigma' & \Sigma\Psi' \\ \Psi\Sigma' & \Psi\Psi'\end{smallmatrix}\right)$ and define the $(m+k) \times (m+k)$ matrix $(\Xi\Xi')^{1/2}$ as $(\Xi\Xi')^{1/2} := (\check{\Sigma}, \check{\Psi})'$. As a result

$$dZ(t) = (a_Z + A_Z\hat{X}(t))\,dt + \check{\Psi}\,dU(t), \quad Z(0) = z,$$

$$dS_i(t) = (a + \tilde{A}\hat{X}(t))_i\,dt + \sum_{k=1}^{M+k}\check{\sigma}_{ik}\,dU_k(t),$$

$$S_i(0) = s_i. \qquad (13)$$

The filtering problem is unrelated to the subsequent stochastic control problem: the dynamics of $\hat{X}(t)$ will be the same for all investors regardless of their risk aversion or time horizon.

2.4. Solve the stochastic control problem

The key is to express and solve a stochastic control problem in which $X(t)$ is replaced by $\hat{X}(t)$ and the dynamic equation (1) by the Kalman filter. Optimal strategies take the form $h(t, \hat{X}(t))$.

Davis and Lleo solve a risk-sensitive asset management problem, where the investor's objective is to maximise the criterion

$$J(t, x, h; T, \theta) = -\frac{1}{\theta}\ln\mathbb{E}[e^{-\theta\ln V_T}]$$
$$= -\frac{1}{\theta}\ln\mathbb{E}[V_T^{-\theta}]. \qquad (14)$$

This criterion relates the evolution of the investor's wealth, $V(t)$, with the investor's risk sensitivity $\theta \in (-1, 0) \cup (0, \infty)$.

The optimal asset allocation $h^*(t)$ for this stochastic control problem is

$$h^*(t) = \frac{1}{1+\theta}(\check{\Sigma}\check{\Sigma}')^{-1}[\hat{a} + \tilde{A}\hat{X}(t) - \theta\check{\Sigma}\check{\Lambda}'(t)D\Phi], \quad (15)$$

where the value function Φ has the form $\tilde{\Phi}(t, x) = \frac{1}{2}x'Q(t)x + x'q(t) + k(t)$. Here, $Q(t)$ satisfies a Riccati equation, $q(t)$ solves a system of linear Ordinary Differential Equations and $k(t)$ can be calculated by direct integration.

3. Including predictive models in the Black–Litterman model in continuous time

Most analysts formulate their views and confidence intervals based on the output of macroeconomic models or of valuation models. We discuss how to incorporate the output of predictive models in the Black–Litterman model in continuous time. The general idea is to treat the output of a predictive model as a views, and to model it with a stochastic differential equation of the same form as Equation (5). We use the model's prediction to calibrate the drift of the view, and the model's error to estimate the diffusion term. Because the Kalman filter gives more weight to more credible observations and less weight to less credible observations, more accurate predictive models will have more influence on the asset allocation.

We select specific examples of predictive models: the CAPE and the BSEYD.

3.1. The Campbell–Shiller model and the CAPE

Campbell and Shiller (1988) initially proposed a vector-autoregressive model relating the log return on the S&P 500 with the log dividend-price ratio, lagged dividend growth rate and average annual earnings over the previous 30 years. They performed a regression of the log returns on the S&P 500 over 1 year, 3 years and 10 years against each of these variables and over the average annual earnings

over the previous 10 years. Using average earnings rather than current earnings to compute the price-to-earnings (P/E) ratio reduces its sensitivity to current economic and market conditions.[†]

To perform the regression, Campbell and Shiller defined the one-period total return on the stock as $h_{1t}^{\text{beg}} :=$ $\ln((P_{t+1}^{\text{beg}} + D_t)/P_t^{\text{beg}})$, where P_t^{beg} is the level of the S&P 500 at the beginning of period t and D_t is the dividend received during period t. The i period total return on the stock is $h_{it}^{\text{beg}} := \sum_{j=0}^{i-1} h_{1,t+j}$. The regression is

$$h_{it}^{\text{beg}} = a + b \ln \left(\frac{P_t^{\text{beg}}}{\overline{E_{t,-n}^{\text{beg}}}} \right) + \epsilon_t, \qquad (16)$$

where P_t is the level of the S&P 500 index at time t and $\overline{E_{t,-n}^{\text{beg}}}$ is the average of past annual earnings over the last n years, namely $\overline{E_{t,-n}^{\text{beg}}} = (1/n) \sum_{i=0}^{n-1} E_{t-i}^{\text{beg}}$. The R^2 computed by Campbell and Shiller (1988) for $n = 30$ is 0.566, which is higher than the 0.401 computed for $n = 10$ and higher than the R^2 of regressions against the log dividend-price ratio and lagged dividend growth rate. Hence, the regression based on 30-year average earnings has a greater explanatory power than the regression based on 10 years of earnings.

The CAPE is a direct descendant of this research. It is the P/E ratio computed using 10-year average earnings (see Shiller 2015).

We can use the CAPE to create a view about the future evolution of the equity risk premium. Using historical data, we perform a regression of the equity risk premium at an horizon h against the logarithm of the CAPE

$$y_t^h = a + bx_t^n + \epsilon_t, \qquad (17)$$

where $y_t^h = \ln(\tilde{P}_{t,h}/\tilde{P}_t)$ is the risk premium at a horizon of h years; $x_t^{10} = \ln(P_t/\overline{E}_{t,-10})$ is the logarithm of the CAPE; $\tilde{P}_t = P(t)/S_0(t)$ and $\tilde{P}_{t,h} = P(t+h)/S_0(t+h)$ are, respectively, the discounted value of the S&P 500 at time t and $t + h$.

For clarity, we have dropped the $^{\text{beg}}$ superscript and consider the S&P 500 with all dividends reinvested (S&P 500 Total Return Index).

By varying the time horizon h from 1 year to 10 years, we can use the regressions to predict the evolution of the equity risk premium at various points over a 10-year horizon. The point estimates for $h = 1, \dots, 10$ provide the data to fit the functions $a_Z(t)$ and $A_Z(t)$ in the view process (5). Then we use the distribution of the error term ϵ_t to fit the diffusion term $\Psi(t)$.

3.2. Crash prediction models and the BSEYD

Prediction models for equity market crashes generate a signal to indicate a downturn in the equity market at a given horizon h. Example of crash prediction models include the BSEYD discussed in Ziemba and Schwartz (1991), the high P/E model (Lleo and Ziemba in press), the variations on Warren Buffett's market value-to-the-Gross National Product measure (Lleo and Ziemba 2015a) or the continuous time disorder detection model (Shiryaev *et al.* 2014, 2015).

The signal occurs whenever the value of a crash measure crosses a threshold. Given a prediction measure $M(t)$, a crash signal occurs whenever

$$\text{SIGNAL}(t) = M(t) - K(t) > 0, \qquad (18)$$

where $K(t)$ is a time-varying threshold for the signal. Three key parameters define the signal: (i) the choice of measure $M(t)$; (ii) the definition of threshold $K(t)$ and (iii) the specification of a time interval H between the occurrence of the signal and that of an equity market downturn.

The BSEYD relates the yield on the S&P 500 (measured by the earnings yield, which is the reciprocal of the P/E ratio) to the yield on nominal Treasury bonds

$$\text{BSEYD}(t) = r(t) - \rho(t) = r(t) - \frac{E(t)}{P(t)}, \qquad (19)$$

where $\rho(t)$ is the earnings yield at time t and $r(t)$ is the most liquid (10- or 30-year) Treasury bond rate $r(t)$. The BSEYD was initially developed for the Japanese market shortly before the crash of 1990 (Ziemba and Schwartz 1991), and it has since been used successfully on a number of international markets (Lleo and Ziemba 2015b).

The threshold K depends on a confidence interval, calculated using a moving average and standard deviation of the BSEYD measure. Usually, the one-tailed confidence interval is established at a 95% level. This corresponds to 1.645 standard deviations above the mean for a normal distribution.

The time horizon is generally set to one year (252 trading days).

Lleo and Ziemba (in press) tested the accuracy of 32 model specifications of the P/E ratio and BSEYD measures over more than 50 years, from 1962 to 2014. Over this period, the S&P 500 experienced 21 downturns, defined as declines of at least 10% peak-to-trough over a maximum of one year. On average, these downturns lasted for 265 days and resulted in a 21.87% decline in the index. The authors found that the BSEYD produced 38 signals, of which 29 signals were followed by an equity market downturn. At 76.32%, the BSEYD's accuracy is statistically significant. Figure 1 displays the cumulative return on the S&P 500 for the two years following a crash signal.

To illustrate the procedure, we consider the case where the BSEYD is currently producing a crash signal. Here,

---

[†]Graham and Dodd (1934) already warned against this shortcoming and advocated the use of a P/E ratio based on average earnings over 10 years.
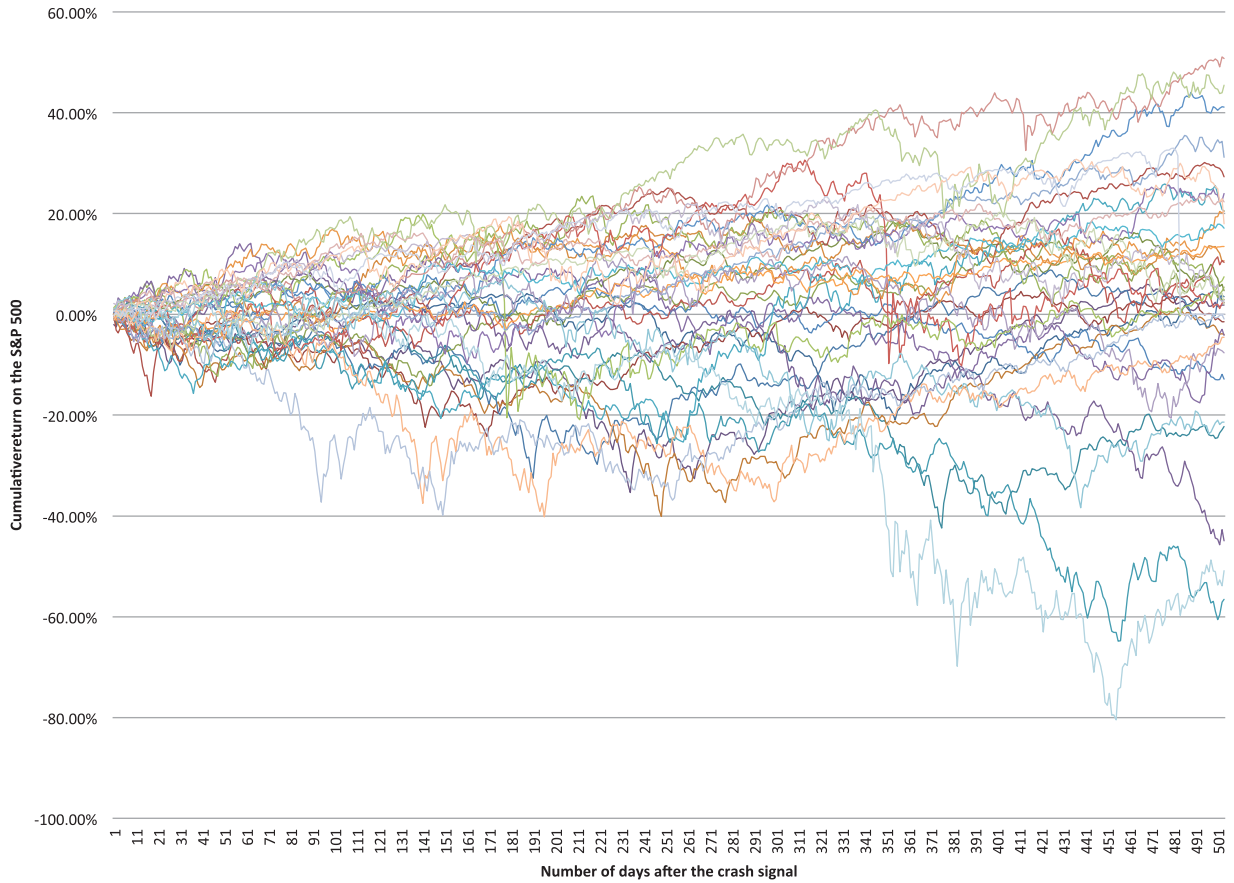
Figure 1.   Cumulative return on the S&P 500 after a crash signal.

we model the view with a generalised Ornstein–Uhlenbeck process (see Hull and White 1994)

$$dZ(t) = (a_Z(t) - A_Z Z(t)) \, dt + \Psi \, dW_{n+m+1}(t). \quad (20)$$

The solution to this SDE is

$$Z(t) = z \, e^{-A_Z t} + \int_0^t e^{-A_Z(t-s)} a_Z(s) \, ds$$
$$+ \Psi \int_0^t e^{-A_Z(t-s)} \Psi \, dW_{n+m+1}(s), \quad (21)$$

with

$$\mathbb{E}[Z(t)] = z \, e^{-A_Z t} + \int_0^t e^{-A_Z(t-s)} a_Z(s) \, ds,$$

$$\mathrm{Var}[Z(t)] = \frac{\Psi^2}{2A_Z}(1 - e^{-2A_Z t}). \quad (22)$$

This process gives us enough flexibility to incorporate crash predictions, while keeping the affine form necessary for an efficient resolution of the filtering and optimisation problems.

A crash prediction model produces a binary signal: either it predicts a crash or it does not. As a result, we cannot map a crash prediction directly into a view. We need to turn the crash prediction into a return prediction by looking at the evolution of the risk premium conditional on the crash signal. The signals produced by the BSEYD provide us with 38 paths for the risk premium conditional on a

crash signal. We can use these paths to calibrate the parameters of the view process by matching the moments of the stochastic process in Equation (22) with the empirical moments inferred from these historical paths.

A typical downturn lasts for slightly more than one calendar year. Out of the 21 downturns, only 4 downturns lasted more than two years, and none lasted more than two years and a half. To capture the evolution of the risk premium during these downturns, we need to consider at least a two-year time horizon from the signal. On the other hand, we have 38 signals in 50 years; so we cannot have a time horizon of more than a couple of years without having a risk that the signals will interfere with each other. This leads us to concentrate on the evolution of the risk premium (conditional on a crash signal) over two years exactly.

If the horizon t of the optimisation is longer, say $T = 5$ years, we will need to make an assumption on the behaviour of the risk premium between the two-year horizon of the crash prediction model and the five-year horizon of the optimisation. For simplicity, we assume in this paper that the risk premium converges linearity to a long-term average of 4%.

We calibrate the mean evolution of the view process to the mean path of the risk premium conditional on a crash signal. To make the calibration easier, we look for a polynomial function. This choice is convenient because polynomial functions have the advantage of being smooth and bounded over a bounded domain. They can also be
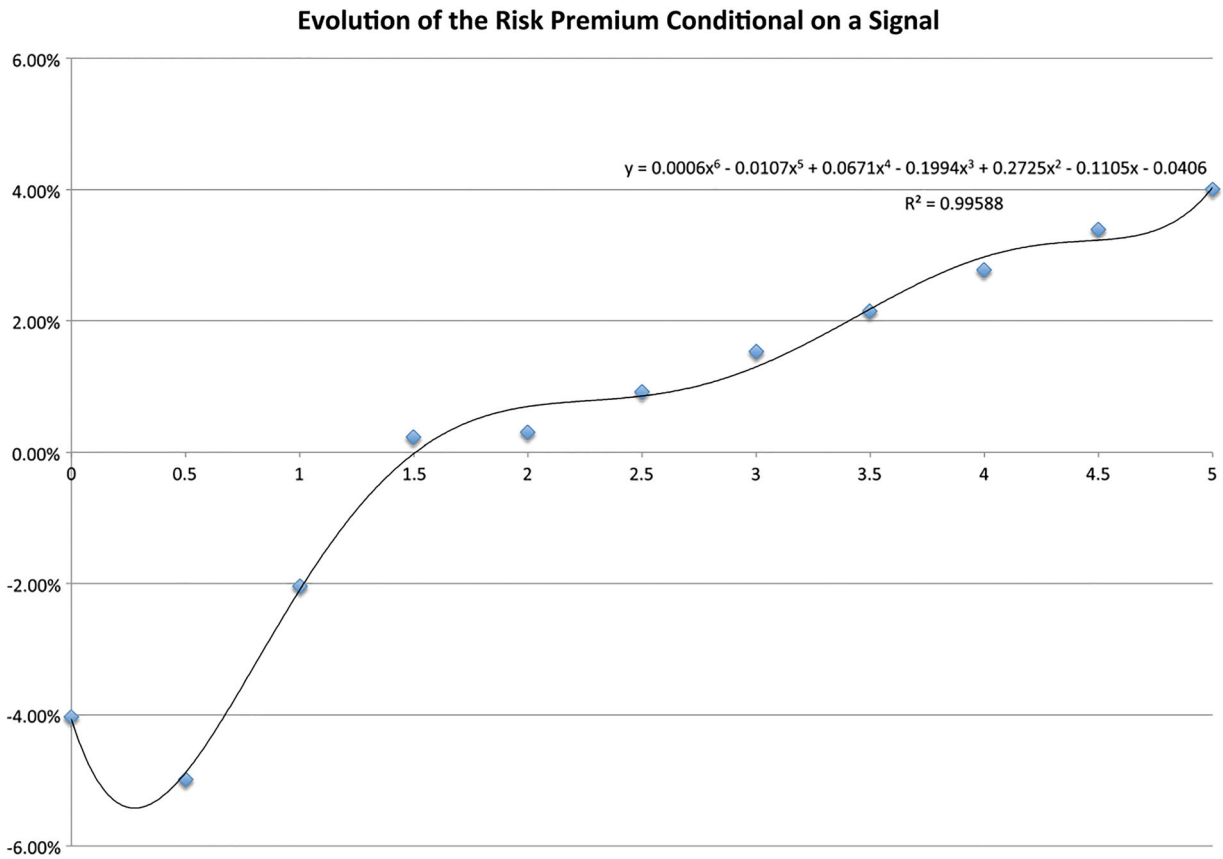
Figure 2.   Polynomial calibration function.

viewed as a Taylor expansion of the true functional relationship between the view and time. Figure 2 suggests that the sixth-order polynomial function

$$\mathcal{P}(t) = 0.0006t^6 - 0.0107t^5 + 0.0671t^4 - 0.1994t^3$$
$$+ 0.2725t^2 - 0.1105t - 0.0406$$

provides a close fit to the actual data. To fit this polynomial function, we express the function a_Z as a fifth-order polynomial

$$a_Z(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4. \qquad (23)$$

Selecting $A_Z = \frac{1}{2}\ln 2 = 0.3466$ implies a half life of two years. To finish the calibration, we perform a Taylor expansion of $\mathbb{E}[Z(t)]$ around $t=0$. Matching the terms of the Taylor expansion with the polynomial given in Equation (23), we get $z = -4.06\%, \beta_0 = -0.1246, \beta_1 = 0.5067, \beta_2 = -0.5038, \beta_3 = 0.1993, \beta_4 = -0.0302$ and $\beta_5 = 0.0006$.

To get the diffusion parameter Ψ, we match the variance of $Z(t)$ to the highest annualised variance across the 38 historical paths, equal to 29.74%. From Equation (22), we get $\Psi = 29.74\% \times \sqrt{2A_Z/1 - e^{2A_Z}} = 31.82\%$.

4.   Conclusion

The ability to include predictive models alongside analyst views and historical data addresses one of the main challenges of stochastic optimisation: how to incorporate scenarios while retaining a tractable solution. It also contributes to significantly lowering the downside risk of portfolios and increase their long-term risk-adjusted returns. In this paper, we showed how to incorporate return prediction and crash prediction models as views into the Black–Litterman model in continuous time. Including return prediction models, such as the CAPE is straightforward: we can use the point estimates and model error directly to calibrate the view. Crash prediction models require an extra step: they need to be converted into return prediction models first.

References

Berge, K., Consigli, G. and Ziemba, W.T., The predictive ability of the bond-stock earnings yield differential. *J. Portfolio Manage.*, 2008, **34**, 63–80.

Black, F. and Litterman, R., *Asset Allocation: Combining Investor Views with Market Equilibrium*, 1990, Technical Report (Goldman, Sachs & Co.: New York)

Black, F. and Litterman, R., Global portfolio optimization. *J. Fixed Income*, 1991, **2**, 7–18.

Black, F. and Litterman, R., Global portfolio optimization. *Financ. Anal. J.*, 1992, **48**, 28–43.

Campbell, J.Y. and Shiller, R.J., Stock prices, earnings, and expected dividends. *J. Financ.*, 1988, **43**, 661–676. (Papers and Proceedings of the Forty-Seventh Annual Meeting of the American Finance Association, Chicago, IL, 28–30 December 1987).

Davis, M., Guo, X. and Wu, G., Impulse control of multi-dimensional jump diffusions. *SIAM J. Control Optim.*, 2010, **48**, 5276–5293.

Davis, M. and Lleo, S., Black–Litterman in continuous time: the case for filtering. *Quant. Financ. Lett.*, 2013a, **1**, 30–35.

Davis, M. and Lleo, S., Jump-diffusion risk-sensitive asset management II: jump-diffusion factor model. *SIAM J. Control Optim.*, 2013b, **51**, 1441–1480.

Davis, M. and Lleo, S., Expert Opinions and Behavioral Biases in Continuous Time Asset Allocation, 2015. Available online at: http://ssrn.com/abstract = 2663650 (accessed 20 December 2015).

Davis, M. and Lleo, S., A simple procedure for merging expert opinions to achieve superior portfolio performance. *J. Portfolio Manage*, **42**, in press.

Frey, R., Gabih, A. and Wunderlich, R., Portfolio optimization under partial information with expert opinions. *Int. J. Theor. Appl. Financ.*, 2012, **15**, 1250009.

Gabih, A., Frey, R. and Wunderlich, R., Dynamic programming equations for portfolio optimization under partial information with expert opinions [online], 2014a (accessed 1 December 2014). http://arxiv.org/abs/1303.2513

Gabih, A., Kondakji, H., Sass, J. and Wunderlich, R., Expert opinions and logarithmic utility maximization in a market with Gaussian drift. *Commun. Stoch. Anal.*, 2014b, **8**, 27–47.

Graham, B. and Dodd, D.L., *Security Analysis*, 1934 (Whittlesey House: York, PA).

Hull, J. and White, A., Branhcing out. *Risk*, 1994, **7**, 34–37.

Lleo, S. and Ziemba, W., Does the Bond-stock Earnings Yield Differential Model Predict Equity Market Corrections Better than High P/E Models?, *Financ. Mark. Inst. Instrum.*, in press.

Lleo, S. and Ziemba, W., *Can Warren Buffett also Predict Equity Market Downturns?*, 2015a, Working paper. Available online at: http://ssrn.com/abstract = 2630068.

Lleo, S. and W.T. Ziemba, Some historical perspectives on theBond-Stock Earnings Yield Model for crash prediction around the world. *Int. J. Forecast.*, 2015b, **31**, 399–425.

Shiller, R.J., *Irrational Exuberance*, 2015 (Princeton University Press:Princeton, NJ).

Shiryaev, A.N., Zitlukhin, M. and Ziemba, W.T., When to sell apple and the Nasdaq 100? Trading bubbles with a stochastic disorder model. *J. Portfolio Manage.*, 2014, **40**, 1–10.

Shiryaev, A.N., Zitlukhin, M. and Ziemba, W.T., Land and stock bubbles, crashes and exit strategies in Japan, Circa 1990 and in 2013. *Quant. Financ.*, **15**, 1449–1469.

Ziemba, W. and Schwartz, S., *Invest Japan*, 1991 (Probus: Chicago, IL).