A demystification of the Black–Litterman model: Managing quantitative and traditional portfolio construction

Stephen Satchell*

is a fellow of Trinity College, Cambridge, a lecturer in Economics at Cambridge University and a Visiting Professor at City University Business School, London. His specialities are finance and econometrics, on which he has written at least 80 papers. An academic advisor and consultant to a wide range of financial institutions covering areas such as actuarial valuation, asset management, risk management and strategy design, Dr Satchell has links with research organisations such as the Department of Applied Economics (Cambridge) and INQUIRE, and has doctorates from the Universities of Cambridge and London.

Alan Scowcroft

was educated at Ruskin College, Oxford and Wolfson College, Cambridge where he was awarded the Jennings prize for academic achievement. He taught econometrics at Clare College, Cambridge before joining Phillips and Drew as an econometrician in 1984. There he worked with the leading macro research group of the time building macro-econometric forecasting models and developing innovative software for solving large scale dynamic input-output models. In 1991 he established the equities quantitative research group at UBS with Alun Jones and since that time has worked on every aspect of quantitative modelling from stock valuation to TAA. He has been closely associated with the pioneering work on equity style and portfolio analysis developed by UBS Warburg. Mr Scowcroft's research interests include practical applications of Bayesian econometrics and portfolio optimisation. He is currently the head of equities quantitative research at UBS Warburg.

*Trinity College, Cambridge University, Trinity Lane, Cambridge CB2 1TQ, UK
Tel: +44(0) 1223 338409; Fax: +44(0) 1223 335475; e-mail: ses11@econ.cam.ac.uk

Abstract The purpose of this paper is to present details of Bayesian portfolio construction procedures which have become known in the asset management industry as Black–Litterman models. We explain their construction, present some extensions and argue that these models are valuable tools for financial management.

Keywords: *Bayesian portfolio construction; Black–Litterman model; optimisation; asset allocation*

Introduction

One of the major difficulties in financial management is trying to integrate quantitative and traditional management into a joint framework. Typically, traditional fund managers are resistant to quantitative management, as they feel that techniques of mean-variance analysis and related procedures do not capture effectively their value added. Quantitative managers often regard their judgmental colleagues as idiot savants. Senior management is rarely prepared to intervene when managers are successful and profitable, however they made their decisions. These disharmonies can make company-wide risk-management and portfolio analysis non-operational and can have deleterious effects on company profitability and staff morale.

One model which has the potential to be used to integrate these diverse

approaches is the Black–Litterman (BL) model (Black and Litterman, 1991, 1992). This is based on a Bayesian methodology which effectively updates currently held opinions with data to form new opinions. This framework allows the judgmental managers to give their views/forecasts, these views are added to the quantitative model and the final forecasts reflect a blend of both viewpoints. A lucid discussion of the model appears in Lee (1999).

Given the importance of this model, however, there appears to be no readable description of the mathematics underlying it. The purpose of this paper is to present such a description. In the second and third sections we describe the workings of the model and present some examples. In the fourth section we present an alternative formulation which takes into account prior beliefs on volatility. In the second and third sections, particular attention is paid to the interesting issue of how to connect the subjective views of our managers into information usable in the model. This is not a trivial matter and lies at the heart of Bayesian analysis. Indeed Rev. Bayes had such misgivings about applying Bayes theorem to real-world phenomena that he did not publish his paper (Bayes, 1763): it was presented to the Royal Society by his literary executor (Bernstein, 1996: 129–34).

Workings of the model

Before we present Bayes' theorem and its application by BL to asset management problems, we shall present our notation and basic concepts. We assume that there is an $(n \times 1)$ vector of asset returns $\mathbf{r}$; these are, typically, excess returns measured in the domestic currency and subtracting the domestic cash return which is not included in the vectors. With this convention the asset returns have a well-defined n-dimensional covariance matrix Σ; in particular, their covariance matrix is non-singular. If the returns for period t are denoted by $\mathbf{r}_t$, we shall write $E(\mathbf{r})$ to mean expected forecasted returns. This is shorthand for $E(\mathbf{r}_{t+1} | \mathfrak{I}_t)$, where $\mathfrak{I}_t$ refers to all information up to and including period t. A second related concept is the $(n \times 1)$ vector $\boldsymbol{\pi}$ representing equilibrium excess returns, either in terms of a theory such as the capital asset pricing model (CAPM) or in the sense of the prevailing supply of value-weighted assets. The latter interpretation corresponds to a global market portfolio demonetised in domestic currency.

Algebraically, assuming the validity of the CAPM,

$$\boldsymbol{\pi} = \beta(\mu_{\mathrm{m}} - r_{\mathrm{f}})$$

where μ_{m} is the return on the global market in domestic currency, r_{f} is the riskless (cash) domestic rate of return, β is an $(n \times 1)$ vector of asset betas, where

$$\beta = \mathrm{Cov}(\mathbf{r}, \mathbf{r}' \mathbf{w}_{\mathrm{m}})/\sigma_{\mathrm{m}}^2$$

where $\mathbf{r}' \mathbf{w}_{\mathrm{m}}$ is the return on the global market, $\mathbf{w}_{\mathrm{m}}$ are the weights on the global market, determined by market values, and σ_{m}^2 is the variance of the rate of return on the world market.

If we let $\Sigma = \mathrm{Cov}(\mathbf{r}, \mathbf{r}')$ be the covariance matrix of the n asset classes, then

$$\boldsymbol{\pi} = \delta \Sigma \mathbf{w}_{\mathrm{m}}$$

where $\delta = (\mu_{\mathrm{m}} - r_{\mathrm{f}})/\sigma_{\mathrm{m}}^2$ is a positive constant. If returns were arithmetic with no reinvestment, δ would be invariant to time, since both numerator and denominator would be linear in time. However, if compounding is present, there may be some time effect.

In this paper, we shall only consider

(global) equity in our *n* assets. Extending the model to domestic and foreign equities and bonds presents few difficulties.

Considering Foreign Exchange (*FX*) as an additional asset class does present difficulties as we need to 'convert' currencies into a domestic value, which requires making assumptions about hedge ratios. Black (1990) proves that, in an international CAPM (ICAPM)[1] under very stringent assumptions, all investors hedge the same proportion of overseas investment, and uses this result to justify a global or universal hedging factor which is the same for all investors facing all currencies. Adler and Prasad (1992) discuss Black's result and show how restrictive the result actually is.

It is natural to think of π as being the implied returns from the equilibrium model and, as the above discussion shows, these would depend upon our data and represent the input of the quantitative manager. How can we represent the views of the fund managers? To answer this question, consider Bayes' theorem. In the notation we have defined above, Bayes' theorem states that

$$\text{pdf}(E(\mathbf{r})\,|\,\boldsymbol{\pi}) = \frac{\text{pdf}(\boldsymbol{\pi}\,|\,E(\mathbf{r}))\text{pdf}(E(\mathbf{r}))}{\text{pdf}(\boldsymbol{\pi})}$$

where pdf(.) means probability density function. The above terms have the following interpretations:

— pdf($E(\mathbf{r})$) is the prior pdf that expresses the (prior) views of the fund manager/investor
— pdf($\boldsymbol{\pi}$) represents the marginal pdf of equilibrium returns. In the treatment that follows, it is not modelled. As we will demonstrate, it disappears in the constant of integration.
— pdf($\boldsymbol{\pi}\,|\,E(\mathbf{r})$) is the conditional pdf of

the data equilibrium return, given the forecasts held by the investor.

The result of the theorem pdf($E(\mathbf{r})\,|\,\boldsymbol{\pi}$) is the 'combined' return or posterior forecast given the equilibrium information. It represents the forecasts of the manager/investor after updating for the information from the quantitative model.

The contribution of BL was to place this problem into a tractable form with a prior distribution that was both sensible and communicable to investors. Bayesian analysis has, historically, been weakened by difficulties in matching tractable mathematical distributions to individual's views.

We now review and extend BL's results. We make the following assumptions:

A1 pdf($E(\mathbf{r})$) is represented in the following way. The investor has a set of *k* beliefs represented as linear relationships. More formally, we know the ($k \times n$) matrix **P** and a known ($k \times 1$) vector **q**. Let $\gamma = \mathbf{P}E(\mathbf{r})$ be a ($k \times 1$) vector. It is assumed that $\mathbf{y} \sim N(\mathbf{q}, \Omega)$, where Ω is a ($k \times k$) diagonal matrix with diagonal elements ω_{ii}. A larger ω_{ii} represents a larger degree of disbelief in the relationship represented by γ_i, $\omega_{ii} = 0$ represents absolute certainty, and, as a practical matter, we bound ω_{ii} above zero. The parameters **q** and Ω are called by Bayesians *hyperparameters*; they parameterise the prior pdf and are known to the investor.
A2 pdf($\boldsymbol{\pi}\,|\,E(\mathbf{r})$) is assumed to be $N(E(\mathbf{r}), \tau\Sigma)$ where Σ is the covariance matrix of excess returns and τ is a (known) scaling factor often set to 1. This assumption means that the equilibrium excess returns conditional upon the

individual's forecasts equals the individual's forecast on average. This may not hold in practice as the authors have met many practitioners who have exhibited the most alarming biases relative to the market view. The conditioning needs to be understood in the sense that, if all individuals hold this view and invest in a CAPM-type world, then π represents the equilibrium returns conditional upon the individuals' common beliefs.[2]

Given A1 and A2, it is a straightforward result to show the following theorem.

Theorem 1. The pdf of $E(\mathbf{r})$ given π is given by

$$\text{pdf}(E(\mathbf{r})\,|\,\pi) \sim N([(\tau\Sigma)^{-1} + \mathbf{P}'\,\Omega^{-1}\mathbf{P}]^{-1} \\ [(\tau\Sigma)^{-1}\,\pi + \mathbf{P}'\,\Omega^{-1}\mathbf{q}], \\ [(\tau\Sigma)^{-1} + \mathbf{P}'\,\Omega^{-1}\mathbf{P}]^{-1}]$$

Proof: See Appendix.
We emphasise that Theorem 1 is a result known to Bayesian econometricians and to BL, although they did not report the variance formula in the papers. Also, our interpretation of what is prior and what is sample information may differ from BL.

It should be clear from the previous analysis that neither A1 nor A2 are essential for the model to be used. Most priors used in finance, however, tend to convey little information about the investors' beliefs. Various alternatives such as a diffuse prior (see Harvey and Zhou (HZ), 1990: Equation 6; or Klein and Bawa (KB) 1976: Equation 3) or the more detailed priors presented in Hamilton (1994: Chap. 12) cannot be easily understood in behavioural terms. In Bayesian terms, the prior chosen by BL is called the natural conjugate prior.

Extensions could be considered for volatilities as well. The natural *equilibrium* value for volatility is the Black–Scholes (BS) model, so that if option data were available, the prior on volatility could be updated by the *observed* implied volatility. Unfortunately, the *pdf* of implied volatility would depend on the nature of the stochastic volatility ignored by the BS formula, and there appears to be no simple way forward. An alternative would be to formulate a prior on τ. Although we have no obvious data to update our beliefs, a solution similar to Proposition 12.3 in Hamilton (1994) could be attempted. We present details in the fourth section.

Examples

In this section, we consider various examples which illustrate the methodology.

Example 1

In this example, we consider the case where a sterling-based investor believes that the Swiss equity market will outperform the German by 0.5 per cent per annum. All returns are measured in sterling and are unhedged. This is a modest target and is intended to emphasise that the forecast represents a new *equilibrium* and not short-term outperformance. In the notation of the second section, we have one belief, $k = 1$ in A1. Using the universe of 11 European equity markets listed in Table 2, $\mathbf{P}$ is a (1×11) vector of the form

$$\mathbf{P} = [1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0]$$

and $\mathbf{q} = 0.5$ per cent. Table 1 lists the parameters used to compute the conditional forecast.

The computed values $E(\mathbf{r}|\pi)$ are shown in Table 2. In addition to the prior view of the relative performance of Swiss and

Table 1 Bayesian parameters

Parameter	Value	Symbol
Delta	5.00	δ
Tau	1.00	τ
View	0.05	q
Confidence	0.05	ω

Table 3 Optimisation parameters

Parameter	Value
Risk aversion λ	2.5
Tracking error limit	2.5
Portfolio beta	1.0

German markets, larger changes from the implied view for other markets are associated with low covariance with the Swiss market. In Table 3, we report certain key parameters associated with our portfolio construction.

We now consider the impact of the conditional forecast in an optimisation problem, where the objective is a simple mean–variance utility function. The risk–aversion parameter has been set with reference to delta. The beta of the portfolio and the sum of the weights are constrained to unity. The results presented in Table 4, show, as expected, a switch from the German to the Swiss market. Some large differences in forecasts, Italy for example, are translated into small changes in the portfolio weights as the optimiser takes into account the benchmark weight, the asset beta and the impact of covariances.

In both this and the following example, currency holdings were free to vary between zero and minus the market

weight (ie from unhedged to fully hedged). The assumed benchmark holding of currency is zero for all markets. The solution weights for currencies, which are not shown in the table, are all negligible.

The Sharpe ratio for the solution is 0.16 with a tracking error[3] of 0.39, the portfolio is beta constrained to 1.0.

Example 2

In this second example, we consider the case where a US dollar-based investor believes that six *hard currency* markets will outperform nine other European markets, on average, by 1.5 per cent per annum. This could be interpreted as a possible EMU scenario. As in Example 1, this still represents one view and **P** is now a (1×15) vector equal to

$$[1/6 \ldots 1/6 - 1/9 \ldots -1/9]$$

The values for the other parameters are as shown in Table 5. Note that Delta (δ)

Table 2 Forecast results

Market	Bench weight	Swiss Cov × 100	π	$E(r|\pi)$	Difference
Switzerland	0.0982	0.1884	5.34	5.53	+0.19
Germany	0.1511	0.0724	6.46	6.31	−0.15
Denmark	0.0136	0.0766	5.31	5.29	−0.02
Spain	0.0409	0.0666	8.07	7.99	−0.08
Finland	0.0125	0.0666	10.69	10.55	−0.14
France	0.1234	0.1016	7.93	7.89	−0.03
Italy	0.0568	0.0061	8.06	7.88	−0.18
Netherlands	0.0870	0.0826	5.64	5.62	−0.03
Norway	0.0103	0.0979	8.43	8.40	−0.03
Sweden	0.0474	0.0776	7.71	7.67	−0.04
UK	0.3588	0.0784	6.33	6.33	−0.00

Table 4 Optimisation results

Market	Beta	Benchmark weight (%)	Solution weight (%)	Difference
Switzerland	0.80	9.82	12.19	+2.37
Germany	0.97	15.11	12.81	−2.30
Denmark	0.80	1.36	1.22	−0.14
Spain	1.20	4.09	4.27	+0.18
Finland	1.57	1.25	1.37	+0.11
France	1.18	12.34	12.61	+0.27
Italy	1.20	5.68	5.63	−0.05
Netherlands	0.85	8.70	8.04	−0.66
Norway	1.25	1.03	1.10	+0.07
Sweden	1.15	4.74	4.77	+0.02
UK	0.95	35.88	36.00	+0.12

Table 5 Bayesian parameters

Parameter	Value	Symbol
Delta	3.000	δ
Tau	1.000	τ
View	0.015	q
Confidence	0.025	ω

has now been set at 3 to ensure that the level of the conditional forecast accords with historical experience.

The conditional forecast is shown in Table 6. The difference between the implied and conditional forecasts is broadly in line with the imposed view, with the exception that the forecast for Ireland actually goes down while Switzerland increases slightly.

To consider the impact of the conditional forecast, we solve a simple optimisation problem, where, as in Example 1, the asset weights are constrained to be positive and sum to unity. The asset beta is constrained to unity and currency weights are free to vary between unhedged and fully hedged for each market. The tracking error is bounded at 2.5 (see Table 7).

The optimisation results are shown

Table 6 Forecast results

	Bench weight	π	$E(r|\pi)$	Difference
'Hard' markets				
Austria	0.0060	14.84	15.05	+0.21
Belgium	0.0244	13.75	13.83	+0.08
France	0.1181	14.86	14.98	+0.12
Germany	0.1446	13.57	13.60	+0.04
Netherlands	0.0832	12.33	12.38	+0.06
Ireland	0.0076	11.18	11.03	−0.15
'Soft' markets				
Denmark	0.0130	12.24	11.92	−0.32
Finland	0.0120	18.83	17.58	−1.24
Italy	0.0543	16.62	15.42	−1.20
Norway	0.0098	15.55	15.04	−0.51
Portugal	0.0049	11.84	11.67	−0.17
Spain	0.0392	13.63	13.03	−0.60
Sweden	0.0454	13.04	12.28	−0.76
Switzerland	0.0940	13.27	13.29	+0.02
UK	0.3433	12.74	12.68	−0.06

Table 7 Optimisation parameters

Parameter	Value
Risk aversion λ	1.5
Tracking error limit	2.5
Portfolio beta	1.0

in Table 8 and not surprisingly show a positive tilt in favour of the *strong currency* markets. Interestingly, even though the optimiser was free to hold currency up to a fully hedged position, all the solution weights for currencies are zero. The Sharpe ratio for the solution is 0.18 with a tracking error of 1.8. The portfolio beta is constrained to unity.

Overall, we feel that the examples justify our confidence in the approach. Care needs to be taken interpreting the conditional forecast, however, since it is the product of the prior view *and* the data model. In these examples, the data model has been taken to be the implied excess returns generated by a mean–variance optimisation problem. Even though such excess returns can be counter-intuitive, as in the case of Ireland in Example 2, they may be understood as the extent to which the neutral forecast has to change to reflect properly the views held by the investor. When these excess returns are subsequently fed back into the optimisation process, the investor's optimal weights will reflect the prior view.

It is this usage of implied excess returns in the data model which also helps to address one of the principal reservations many practitioners have with respect to the use of optimisers in portfolio construction, namely their extreme sensitivity to changes in forecasts. Raw forecast alphas are inevitably volatile and, if used as optimiser inputs, give rise to completely unacceptable revisions to portfolio weights. By combining neutral model forecasts with the investor's views, the Bayesian formulation produces robust inputs for the optimiser.

Alternative formulations

In this section, we present two alternative formulations of the BL model, the first of which takes into account prior beliefs about overall volatility. To do this, we make the following adjustment. We shall assume that τ is now unknown and stochastic so that

Table 8 Optimisation results

Market	Beta	Benchmark weight (%)	Solution weight (%)	Difference
Austria	1.09	0.60	3.27	+2.67
Belgium	1.02	2.44	4.82	+2.38
France	1.10	11.81	15.90	+4.09
Germany	1.00	14.46	18.63	+4.17
Netherlands	0.92	8.32	9.28	+0.96
Ireland	0.84	0.76	3.39	+2.63
Denmark	0.91	1.30	0.00	−1.30
Finland	1.37	1.20	0.00	−1.20
Italy	1.22	5.43	2.66	−2.77
Norway	1.14	0.98	0.00	−0.98
Portugal	0.88	0.49	0.00	−0.49
Spain	1.01	3.92	0.98	−2.94
Sweden	0.97	4.54	1.82	−2.72
Switzerland	0.98	9.40	6.79	−2.61
UK	0.95	34.33	32.46	−1.88

A3

$$\text{pdf}(\boldsymbol{\pi} \,|\, E(\mathbf{r}), \tau) \sim N(E(\mathbf{r}), \tau\Sigma).$$

Furthermore,

$$\text{pdf}(E(\mathbf{r}) \,|\, \tau) \sim N(\mathbf{q}, \tau\boldsymbol{\Omega}),$$

A4 The marginal (prior) pdf of $\omega = 1/\tau$ is given by the following,[4]

$$\text{pdf}(\omega) = \frac{(\lambda/2)^{m/2}\,\omega^{(m/2)-1}\exp\left(-\dfrac{\lambda\omega}{2}\right)}{\Gamma(m/2)},$$

$$0 < \omega < \infty$$

This pdf has two hyperparameters m and λ, and we assume it is independent of pdf($\boldsymbol{\pi}$).

Remark 1. Here we treat τ as a fundamental parameter that measures the overall dispersion of $\boldsymbol{\pi}$ about $E(\mathbf{r})$. Considering $\text{pdf}(E(\mathbf{r}) \,|\, \tau)$, we define the elements of $\boldsymbol{\Omega}$ relative to τ so that $\omega_{ii} = 1$ reflects a degree of disbelief equal in scale to the dispersion measure of $\boldsymbol{\pi}$ about $E(\mathbf{r})$, a value $\omega_{ii} > 1$ implies greater disbelief than before and an increase in τ not only moves the dispersion of equilibrium expected returns about the forecasts but also increases the overall degree of disbelief in the forecasts.

Remark 2. The prior pdf of $\omega = 1/\tau$ is a scale gamma, where ω is often called the precision. It follows that $E(\omega) = m\lambda$ and $\text{Var}(\omega) = 2m\lambda^2$. This means that for fixed $E(\omega)$ as $m \to \infty$, $\text{Var}(\omega) \to 0$ and hence is a more reliable prior.

We are now in a position to state our new result which is, again, a standard result in the Bayesian literature. For a similar result, see Hamilton (1994, Proposition 12.3).

Theorem 2. If we assume A3 and A4, then

$$\text{pdf}(E(\mathbf{r}) \,|\, \boldsymbol{\pi}) \propto$$
$$[m + (E(\mathbf{r}) - \theta)'\,\lambda^\star\, V\,(E(\mathbf{r}) - \theta)]^{-(m+n)/2}$$

which is a multivariate t distribution. The vector θ is the term $E(\mathbf{r} \,|\, \boldsymbol{\pi})$ given in Theorem 1, the matrix V is the $\text{Var}(\mathbf{r} \,|\, \boldsymbol{\pi})$ given in Theorem 1 while

$$\lambda^\star = \frac{m}{\lambda + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C}}$$

where $\mathbf{A}$, $\mathbf{C}$, $\mathbf{H}$ are defined in the proof of Theorem 1.

Proof: See Appendix.

An immediate corollary of Theorem 2 is the following.

Corollary 2.1: $\text{pdf}(\omega \,|\, E(\mathbf{r}), \boldsymbol{\pi})$ is a scale gamma with 'degrees of freedom' $m + n$ and scale factor $\mathbf{G} + \lambda$, where $\mathbf{G} = (\boldsymbol{\pi} - E(\mathbf{r}))'\Sigma^{-1}(\boldsymbol{\pi} - E(\mathbf{r})) + (\mathbf{P}E(\mathbf{r}) - \mathbf{q})'\boldsymbol{\Omega}^{-1}(\mathbf{P}E(\mathbf{r}) - \mathbf{q})$.

Proof: See Appendix.

The consequence of Corollary 2.1 is that we can now compute

$$E(\omega \,|\, E(\mathbf{r}), \boldsymbol{\pi}) = (m + n)(\mathbf{G} + \lambda)$$

and

$$\text{Var}(\omega \,|\, E(\mathbf{r}), \boldsymbol{\pi}) = 2(m + n)(\mathbf{G} + \lambda)^2$$

The increase in precision can be computed as

$$E(\omega \,|\, E(\mathbf{r}), \boldsymbol{\pi}) - E(w)$$
$$= (m + n)\,(\mathbf{G} + \lambda) - m\lambda$$
$$= m\mathbf{G} + n(\mathbf{G} + \lambda)$$

It is interesting to note that, although our expected returns now have a multivariate t distribution, such a returns distribution is consistent with mean–variance analysis and the CAPM.

[This is proved in Klein and Bawa (1976)]. Thus, our extended analysis leaves us with a mean vector and a covariance matrix which, up to a scale factor, are the same as before. What we gain is that probability computations will now involve the use of the t distribution. This will give the same probabilities as the normal for large m, but for small m will put more weights in the tails of our forecast distribution. Thus, we can manipulate this feature to give extra diagnostics to capture uncertainties about our forecasts.

We do not present numerical calculations for this model, as the nature of the prior is too complex to capture the beliefs of a typically non-mathematical fund manager. However, in our experience, fund managers are able to provide a range of scenarios for expected returns and associate probabilities with these scenarios. We shall explore such a model next, this being the second 'extension' of the BL model referred to earlier.

A5 The prior pdf for $E(\mathbf{r})$ is of the form $\mathbf{P}E(\mathbf{r}) = q_i$, $i = 1,\ldots,m$. Each (vector value q_i has prior probability p_i, where $\sum_{i=1}^{m} p_i = 1$. $\mathbf{P}$ and $E(\mathbf{r})$ have the same definition as before.

If we combine A5 with A3, it is straightforward to compute

$$p_i^\star = \mathrm{prob}(\mathbf{P}E(\mathbf{r}) = q_i \,|\, \boldsymbol{\pi}); \quad \sum_{i=1}^{m} p_i^\star = 1$$

$$p_i^\star = \frac{\mathrm{pdf}(\boldsymbol{\pi}\,|\,\mathbf{P}E(\mathbf{r}) = q_i)\,\mathrm{prob}(E(\mathbf{r}) = q_i)}{\mathrm{pdf}(\boldsymbol{\pi})}$$

and

$$\mathrm{pdf}(\boldsymbol{\pi}) = \sum_{i=1}^{m} \mathrm{pdf}(\boldsymbol{\pi}\,|\,\mathbf{P}E(\mathbf{r}) = q_i)\,\mathrm{prob}(E(\mathbf{r}) = q_i)$$

$$\mathrm{pdf}(\boldsymbol{\pi}\,|\,\mathbf{P}E(\mathbf{r}) = q_i) = \varphi_i \tag{1}$$

$$= \left(\frac{1}{2\pi}\right)^{k/2} \frac{1}{\det(\mathbf{P}\Sigma\mathbf{P}')}$$

$$\exp\left(\frac{-(\mathbf{P}\boldsymbol{\pi} - q_i)'\,(\mathbf{P}\Sigma\mathbf{P}')^{-1}\,(\mathbf{P}\boldsymbol{\pi} - q_i)}{2\tau}\right)$$

Combining the above, we deduce that $p_i^\star$ the posterior probability of scenario i becomes

$$p_i^\star = \frac{p_i \varphi_i}{\sum_{i=1}^{m} p_i \varphi_i} \tag{2}$$

Equation (2) gives us an updating rule on the prior probabilities which allows us to rescale our p_i by value of the likelihood function with expected returns evaluated at q_i normalised so that the sum of the weights is one. Thus, if the equilibrium return $\mathbf{p}$ satisfied the $\boldsymbol{\pi}$ condition $\mathbf{P}\boldsymbol{\pi} = q_i$, φ_i would reach its maximum value. We note that since the term in front of the exponential in (1) is common for all φ_i, we can simplify $p_i^\star$ to be

$$\frac{p_i \exp\left(\dfrac{-(\mathbf{P}\boldsymbol{\pi} - q_i)'\,(\mathbf{P}\Sigma\mathbf{P}')^{-1}\,(\mathbf{P}\boldsymbol{\pi} - q_i)}{2\tau}\right)}{\sum_{i=1}^{m} p_i \exp\left(\dfrac{-(\mathbf{P}\boldsymbol{\pi} - q_i)'\,(\mathbf{P}\Sigma\mathbf{P}')^{-1}\,(\mathbf{P}\boldsymbol{\pi} - q_i)}{2\tau}\right)} \tag{3}$$

Our new weights take a maximum value of 1 and a minimum value of 0. Table 9 provides an illustration of the calculations based on ten scenarios for the EMU example given in the third section. Column one shows the assumed outperformance of the strong currency markets for each scenario. For simplicity, we have assumed that the manager believes each scenario to be equally likely; $p_i = 0.1$. The calculated posterior probabilities $p_i^\star$ show clearly that substantial outperformance is much less

Table 9 Posterior probabilities

Scenario outperformance (%)	Prior probability (%)	Posterior probability (%)
0.5	10.00	10.07
1.0	10.00	10.06
1.5	10.00	10.05
2.0	10.00	10.04
2.5	10.00	10.02
3.0	10.00	10.00
3.5	10.00	9.98
4.0	10.00	9.96
4.5	10.00	9.93
5.0	10.00	9.90

likely given the historic covariances between these markets. In this example, each scenario is associated with only one view. If the scenario contained many views, the posterior probability would still relate to the entire scenario and not an individual view.

In practical terms, the judgmental fund manager can use the posterior probability $p_i^\star$ as a consistency check of the prior belief associated with scenario i expressed as probability p_i. If the scenario seems unlikely when tested against the data using (3) the confidence numbers ω_{ii} defined in A1 can be revised upwards accordingly. Equation (3) can therefore be regarded as a useful adjunct to Theorem 1 by helping the rational manager formulate the inputs required in a Bayesian manner. As observed by no less an authority than Harry Markowitz, 'the rational investor is a *Bayesian*' (Markowitz, 1987: 57, italics in original).

Conclusion

We have presented several examples of Bayesian asset allocation portfolio construction models and showed how they combine judgmental and quantitative views. It is our belief that these models are potentially of considerable importance in the management of the investment process in modern financial institutions where both viewpoints are represented. We present an exposition of these models so that readers should be able to apply these methods themselves. We also present several extensions.

Notes

1   Here we use the acronym ICAPM to mean international CAPM. The standard usage for ICAPM is for intertemporal CAPM. Since the international CAPM is a particular application of Merton's intertemporal CAPM, this should cause no confusion.
2   This rather loose interpretation can be tightened; see Hiemstra (1997) for a construction of a CAPM model based on heterogeneous expectations by investors.
3   The tracking error or active risk of a portfolio is conventionally defined as the annualised standard deviation of portfolio active return (ie the excess return attributable to holding portfolio weights different from the benchmark weights).
4   $\Gamma(.)$ is the gamma function, $\Gamma(n) = \int_0^\infty x^{n-1} \exp(-x) dx$.

References

Adler, M. and Prasad, B. (1992) 'On Universal Currency Hedges', *Journal of Financial and Quantitative Analysis*, 27(1), 19–39.

Bayes, T. (1763) 'An Essay Toward Solving a Problem in the Doctrine of Chances', *Philosophical Transactions*, Essay LII, 370–418.

Bernstein, P. L. (1996) *Against the Gods*, John Wiley, Chichester.

Black, F. (1989) 'Universal Hedging: Optimal Currency Risk and Reward in International Equity Portfolios', *Financial Analysts Journal*, July–Aug., 16–22, reprinted in *Financial Analysts Journal*, Jan– Feb., 161–7.

Black, F. (1990) 'Equilibrium Exchange Rate Hedging', *Journal of Finance*, July, 45(3), 899–907

Black, F. and Litterman, R. (1991) 'Global Asset Allocation with Equities, Bonds and Currencies', Goldman Sachs and Co., October.

Black, F. and Litterman, R. (1992) 'Global Portfolio Optimization', *Financial Analysts Journal*, Sept–Oct, 28–43.

Hamilton, J. (1994) *Time Series Analysis*, Breedon University Press, Princeton, NJ.

Harvey, C. R. and Zhou, G. (1990) 'Bayesian Inference in Asset Pricing Tests', *Journal of Financial Economics*, 26, 221–54.

Hiemstra, C. (1997) *Addressing Nonlinear Dynamics in Asset Prices with a Behavioural Capital Asset Pricing Model*, Accounting and Finance Department,

University of Strathclyde.

Klein, R. W. and Bawa, V. S. (1976) 'The Effect of Estimation Risk on Optimal Portfolio Choice', *Journal of Financial Economics*, 3, 215–31.

Lee, W. (1999) 'Advanced Theory and Methodology of Tactical Asset Allocation', unpublished manuscript, available at http://faculty.fuqua.duke.edu/~charvey/Teaching/BA453_2000/lee.pdf

Markowitz, H. (1987) 'Mean-Variance Analysis in Portfolio Choice and Capital Markets, Basil Blackwell, Oxford.

Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics*, John Wiley, New York.

Appendix

Proof of Theorem 1

Using Bayes' theorem and Assumptions A1 and A2, we see that

$$\text{pdf}(E(\mathbf{r})\,|\,\boldsymbol{\pi}) = \frac{k \exp\left(-\frac{1}{2\tau}(\boldsymbol{\pi} - E(\mathbf{r}))'\Sigma^{-1}(\boldsymbol{\pi} - E(\mathbf{r})) - \frac{1}{2}(\mathbf{P}E(\mathbf{r}) - \mathbf{q})'\Omega^{-1}(\mathbf{P}E(\mathbf{r}) - \mathbf{q})\right)}{\text{pdf}(\boldsymbol{\pi})}$$

where k is an appropriate constant.

We next simplify the quadratic term in the exponent.

$$E(\mathbf{r})'(\tau\Sigma)^{-1}E(\mathbf{r}) - 2\boldsymbol{\pi}(\tau\Sigma)^{-1}E(\mathbf{r}) + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi} + E(\mathbf{r})'\mathbf{P}'\Omega^{-1}\mathbf{P}E(\mathbf{r}) - 2\mathbf{q}'\Omega^{-1}\mathbf{P}E(\mathbf{r})$$
$$+ \mathbf{q}'\Omega^{-1}\mathbf{q}$$
$$= E(\mathbf{r})'((\tau\Sigma)^{-1} + \mathbf{P}'\Omega\mathbf{P})((\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P})^{-1}((\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P})E(\mathbf{r}) - 2(\boldsymbol{\pi}'(\tau\Sigma)^{-1}$$
$$+ \mathbf{q}'\Omega^{-1}\mathbf{P})((\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P})^{-1}((\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}E(\mathbf{r})) + \mathbf{q}'\Omega^{-1}\mathbf{q} + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}$$

Let
$$\mathbf{C} = (\tau\Sigma)^{-1}\boldsymbol{\pi} + \mathbf{P}'\Omega^{-1}\mathbf{q}$$
$$\mathbf{H} = (\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P}, \text{ we note that } \mathbf{H} \text{ is symmetrical so } \mathbf{H} = \mathbf{H}'$$
$$\mathbf{A} = \mathbf{q}'\Omega^{-1}\mathbf{q} + \boldsymbol{\pi}'(\tau\Sigma)^{-1}\boldsymbol{\pi}$$

We can rewrite the exponent as equal to

$$E(\mathbf{r})'\mathbf{H}'\mathbf{H}^{-1}\mathbf{H}E(\mathbf{r}) - 2\mathbf{C}'\mathbf{H}^{-1}\mathbf{H}E(\mathbf{r}) + \mathbf{A}$$
$$= (\mathbf{H}E(\mathbf{r}) - \mathbf{C})'\mathbf{H}^{-1}(\mathbf{H}E(\mathbf{r}) - \mathbf{C}) + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C}$$
$$= (E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C})'\mathbf{H}(E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C}) + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C}$$

In terms of $E(\mathbf{r})$, terms such as $\mathbf{A} - \mathbf{C}'\mathbf{H}\mathbf{C}$ disappear into the constant of integration. Thus,

$$\text{pdf}(E(\mathbf{r})\,|\,\boldsymbol{\pi}) \propto \exp\left(-\tfrac{1}{2}(E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C})'\mathbf{H}(E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C})\right) \tag{A1}$$

so that $E(\mathbf{r})\,|\,\boldsymbol{\pi}$ has mean $= \mathbf{H}^{-1}\mathbf{C}$ (A2)

$$= [(\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P}]^{-1}[(\tau\Sigma)^{-1}\boldsymbol{\pi} + \mathbf{P}'\Omega^{-1}\mathbf{q}] \tag{A3}$$

and $\text{Var}(\mathbf{r}\,|\,\boldsymbol{\pi}) = [(\tau\Sigma)^{-1} + \mathbf{P}'\Omega^{-1}\mathbf{P}]^{-1}$ (A4)

Proof of Theorem 2

First

$$\text{pdf}(E(\mathbf{r}),w\,|\,\boldsymbol{\pi}) = \frac{\text{pdf}(\boldsymbol{\pi}\,|\,E(\mathbf{r}),w)\text{pdf}(E(\mathbf{r})\,|\,w)\text{pdf}(w)}{\text{pdf}(\boldsymbol{\pi})} \tag{A5}$$

From Assumption A3, we can write

$$\text{pdf}(\boldsymbol{\pi}\,|\,E(\mathbf{r}),w)\text{pdf}(E(\mathbf{r})\,|\,w) = kw^{n/2}\exp-\left(\frac{w\mathbf{G}}{2}\right) \tag{A6}$$

where $\mathbf{G} = (\boldsymbol{\pi} - \mathrm{E}(\mathbf{r}))'\Sigma^{-1}(\boldsymbol{\pi} - E(\mathbf{r})) + (\mathbf{P}E(\mathbf{r}) - \mathbf{q})'\Omega^{-1}(PE(\mathbf{r}) - \mathbf{q})$

If we now use Assumption 4 and Equation (A6), we see that

$$\text{pdf}(E(\mathbf{r}), w\,|\,\boldsymbol{\pi}) = \frac{k\exp\left(-\dfrac{w}{2}(\mathbf{G} + \lambda)\right)\left(\dfrac{\lambda}{2}\right)^{m/2}w^{(m+n)/2-1}}{\Gamma\left(\dfrac{m}{2}\right)\text{pdf}(\boldsymbol{\pi})} \tag{A7}$$

To compute $\text{pdf}(E(r)\,|\,\boldsymbol{\pi})$, we integrate out w.

Let

$$v = \frac{w}{2}(\mathbf{G} + \lambda), \ w = \frac{2v}{(\mathbf{G} + \lambda)}, \ \mathrm{d}w = \frac{2}{(\mathbf{G} + \lambda)}\mathrm{d}v$$

$$\text{pdf}(E(\mathbf{r})\,|\,\boldsymbol{\pi}) = k'\left(\frac{\lambda}{2}\right)^{m/2}\int_0^\infty \exp(-v)\left(\frac{2v}{\mathbf{G} + \lambda}\right)^{(m+n)/2-1}\left(\frac{2}{\mathbf{G} + \lambda}\right)\mathrm{d}v \tag{A8}$$

then

$$= \frac{k'\left(\dfrac{\lambda}{2}\right)^{m/2}2^{(m+n)/2}\Gamma\left(\dfrac{m + n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)(\mathbf{G} + \lambda)^{(m+n)/2}} \tag{A9}$$

The multivariate t is defined (see Zellner, 1971: 383, B20) for matrices $\boldsymbol{\theta}(l \times 1)$ and $\mathbf{V}(l \times l)$ and positive constant v as

$$\text{pdf}(\mathbf{x}\,|\,\boldsymbol{\theta}, \mathbf{V}, v, l) = \frac{v^{v/2}\Gamma((v + l)/2)|\mathbf{V}|^{1/2}[v + (\mathbf{x} - \boldsymbol{\theta})'\mathbf{V}(\mathbf{x} - \boldsymbol{\theta})]^{(l+v)/2}}{\pi^{l/2}\Gamma(v/2)}$$

If we re-write $\mathbf{G} + \lambda$ in terms of $\mathbf{A}$, $\mathbf{C}$ and $\mathbf{H}$ as defined in the proof of Theorem 1, we see that

$$\mathbf{G} + \lambda = (E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C})'\mathbf{H}(E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C}) + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C} + \lambda$$

$$\propto (E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C})'\frac{m\mathbf{H}}{\lambda + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C}}(E(\mathbf{r}) - \mathbf{H}^{-1}\mathbf{C}) + m \tag{A10}$$

This shows that $\text{pdf}(E(\mathbf{r})|\boldsymbol{\pi})$ is multivariate t,

$$\boldsymbol{\theta} = \mathbf{H}^{-1}\mathbf{C} \text{ (as before)}$$

$$\mathbf{V} = \frac{m}{\lambda + \mathbf{A} - \mathbf{C}'\mathbf{H}^{-1}\mathbf{C}}\mathbf{H}, \ l = n \text{ and } \nu = m.$$

Proof of Corollary 2.1

Factorising $\text{pdf}(E(\mathbf{r}),w|\boldsymbol{\pi}) = \text{pdf}(w|E(\mathbf{r}), \boldsymbol{\pi})\text{pdf}(E(\mathbf{r})|\boldsymbol{\pi})$ gives us

$$\text{pdf}(w|E(\mathbf{r}),\boldsymbol{\pi}) = \frac{\text{pdf}(E(\mathbf{r}),w|\boldsymbol{\pi})}{\text{pdf}(E(\mathbf{r})|\boldsymbol{\pi})}$$

thus

$$\text{pdf}(w|E(\mathbf{r}), \boldsymbol{\pi}) = \frac{\dfrac{k \exp\left(-\dfrac{w}{2}(\mathbf{G} + \lambda)\right)\left(\dfrac{\lambda}{2}\right)^{m/2} w^{(m+n)/2-1}}{\Gamma\left(\dfrac{m}{2}\right)\text{pdf}(\boldsymbol{\pi})}}{\dfrac{k'\left(\dfrac{\lambda}{2}\right)^{m/2} 2^{(m+n)/2}\Gamma\left(\dfrac{m+n}{2}\right)}{(\mathbf{G}+\lambda)^{(m+n)/2}\Gamma\left(\dfrac{m}{2}\right)\text{pdf}(\boldsymbol{\pi})}} \tag{A11}$$

using Equations (A6) and A7).

Simplifying,

$$\text{pdf}(w|E(\mathbf{r}), \boldsymbol{\pi}) = \frac{k'' \exp\left(-\dfrac{w}{2}(\mathbf{G} + \lambda)\right)w^{(m+n)/2-1}\mathbf{G}^{(m+n)/2}}{\Gamma\left(\dfrac{m+n}{2}\right)} \tag{A12}$$