The Black-Litterman Model Explained*

Wing CHEUNG†

February, 2009

Abstract

Active portfolio management is about leveraging information. The Black and Litterman Global Portfolio Optimisation Model (BL) (Black and Litterman, 1992) sets information processing in a Bayesian analytic framework. In this framework, the portfolio manager (PM) needs only produce views and the model translates these into security return forecasts. As a portfolio construction tool, the BL model is appealing both in theory and in practice.

Although there has been no shortage of literature exploring it, the model still appears somehow mysterious and suffers from practical issues. This paper is dedicated to enabling a better understanding of this model, and features: -

- ◇ An economic interpretation;
- ◇ A clarification of the model's assumptions and formulation;
- ◇ Implementation guidance;
- ◇ A dimension-reduction technique to enable large portfolio applications; and
- ◇ A full proof of the main result in the Appendix.

We also provide a checklist of practical issues that we aim to address in our forthcoming articles.

JEL Classification: C10, C11, C61, G11, G14

Keywords: asset allocation, portfolio construction, Bayes' Rule, view blending and shrinkage, CAPM, semi-strong market efficiency, mean-variance optimisation, robustness

---

*This paper reproduces an earlier Nomura publication Cheung (2009b) with revisions. I thank Bob Litterman, Ronny Feiereisen, Alan Hofmeyr, Wuan Luo, Amit Manwani, Attilio Meucci, Mayank Mishra, Stephen Vandermark, and numerous European and US quant fund managers for helpful discussions. All mistakes are mine.

†Corresponding author: Nomura International plc, 25 Bank Street, London E14 5LE, UK. Email: wing.cheung@nomura.com; W.Cheung.02@cantab.net

1 Introduction

Active portfolio management is about leveraging forecasts. As a means of forecast, portfolio managers (PM) or analysts collect information, generate views and seek to convert these views into optimal portfolio holdings. These views may not necessarily be explicit security return predictions, but could be views on relative performance or portfolio strategies[1]. On the other hand, portfolio optimisers do not admit views directly as inputs, but rather expects one explicit return forecast for each security. In order to feed an optimiser, PMs need to translate their views into explicit return forecasts for those view-relevant securities, and are forced to come up with a number (often zero) to represent 'no view'. This practice immediately raises two questions:

(1) What is the appropriate way of translating PM views into explicit return forecasts?

(2) Is it legitimate to use zero return to represent 'no view'?

Regarding the second question, zero-mean return forecasts will be treated by the optimiser relentlessly as views. A typical response of the optimiser will be to use this security to leverage others on which the PM expresses optimism. This easily gives rise to 'unexpected' behaviours (that is, unstable, counter-intuitive or corner solutions). Similar 'erratic' behaviours occur in response to estimation errors in the risk model as well. Yet in these situations, the common use of optimisation constraints does not address the underlying problem but definitely undermines the mean-variance efficiency.

To the first question, the Black-Litterman Global Portfolio Optimisation Model (BL) (Black and Litterman, 1992) provides an elegant answer. The model sets the forecast in a Bayesian analytic framework. In this framework, the PM needs only produce a flexible number of views and the model smoothly translates the views into explicit security return forecasts together with an updated covariance matrix - exactly as what a conventional mean-variance portfolio optimiser expects. If the views arrive in an acceptable form, that is, linear views, this model can fully consume them.

Moreover, the model handles the second question with ease: without view, there are theoretical justifications for taking the market equilibrium returns as the default forecasts. A remarkable feature of this approach is robustness. As the posterior views are a combination of the market and the PM views, PMs have a common layer, the market view, as their starting point. Without views, the best strategy is to stick to the market view. With some views, the portfolio should be tilted to reflect these views combined. Since the market view is always considered, it is less likely to run into unstable or corner solutions. In case the PM holds some strong views that dominate the market view, the model also allows the results to be significantly adjusted towards these views. Rather than erratic, this should be considered expected and intuitive.

Based on these considerations, the BL model is appealing in theory and natural in practice. However, we still have not seen wide applications of it. We attribute this to two main reasons:

- The model deserves further explanation.

[1] Views are often also expressed in terms of fundamental or macroeconomic factors. In a forthcoming paper, we will develop a technique for factor-based allocation.

- Some practical issues may have also frustrated applications, for example, confidence parameter setting, alternative views (for example, factor views, stock-specific views), curse of dimensionality (when applied to large portfolios), prior setting (given equilibrium is an abstract concept), optimiser issues, risk model quality, non-linearity, non-normality issues etc.

Although there has been no shortage of literature exploring either the applications or the frontiers of the model (for example, Meucci, 2005; Jones, Lim and Zangari, 2007; Martellini and Ziemann, 2007; Zhou, 2008 etc.), we have seen few documents explaining it (see for example, Satchell and Scowcroft, 2000; Idzorek, 2004); let alone taking these practical issues seriously. We have therefore done some significant work to make the model practical. As a starting point, this article focuses on an explanation of the original model. By exploring the information processing challenges encountered in a typical portfolio management process, we enrich Black and Litterman (1992)'s original motivation for the BL model. We then establish that the BL model is buttressed by three pillars: the *Semi-Strong Market Efficiency* assumption, the *Capital Asset Pricing Model* (CAPM), and the *Bayes' Rule*. With the assistance of a carefully chosen notation system, we formulate the model with particular attention to its technical details, that is, model assumptions, main results, and a full proof (in the Appendix), to unveil the intrinsic logic. Implementation guidance then follows. In order to enable large portfolio applications, we also discuss a dimension-reduction technique that resolves the high-dimensionality issue before we reach our concluding remarks.

We will address a list of other practical issues in our forthcoming articles as we develop new thoughts around the topic.

Let us examine what motivates the BL model first.

2 Information Processing, Traditional Portfolio Optimisation and Its Weaknesses

Suppose there are n securities in the investment universe. Assuming normality, the distribution of the security returns are fully determined by their first and second moments, that is, $\vec{\tilde{r}}_{[n\times 1]} \sim \mathbb{N}(\vec{m}_{[n\times 1]}, \mathbf{V}_{[n\times n]})$[23], where $\vec{m}$ is the vector of real mean security returns and $\mathbf{V}$ is the real variance-covariance matrix. These moments are not directly observable.

In a typical portfolio management process, people acquire public market information $\mathcal{G}$ together with some private information $\mathcal{H}$ in order to assess $\vec{m}$ and $\mathbf{V}$. Considering the public information first, $\mathcal{G}$ typically includes announced background economic driving information, historical market data, market consensus (and maybe, mis-perception), and announced company-specific news, and so on. The information accrues over time. With only the common market information $\mathcal{G}$, continuing to assume normality, the perceived

[2]In this article, we use upper-case R to stand for total return and lower-case r for excess return, that is, $r = R - r_f$, where r_f is the risk-free rate.

[3]In this article, we use "$\tilde{x}$" to denote a random variable; "$\vec{x}$" to denote a vector; and a bold symbol "$\mathbf{X}$" to denote a matrix. The dimension(s) of vector and matrix will be clarified on its first appearance. For example, $\vec{\tilde{r}}_{[n\times 1]}$ stands for the n by 1 return vector with random entries.

security returns distribution can be represented by the estimated first and second moments, that is, $\vec{\tilde{r}}_{|\mathcal{G}} \sim \mathbb{N}(\vec{\widehat{\mu}}_{[n\times 1]}, \mathbf{\Sigma}_{[n\times n]})$[4], where $\vec{\widehat{\mu}} = \mathbb{E}(\vec{\tilde{r}}|\mathcal{G})$[5] is the vector of mean estimates and $\mathbf{\Sigma} = \mathbb{E}(\mathbf{V}|\mathcal{G})$ is the estimated covariance matrix. The second-moment estimate $\mathbf{\Sigma}$ is generally regarded as more reliable than the first-moment estimates $\vec{\widehat{\mu}}$. The latter is the *holy grail* of the investment industry.

On the other hand, the private information $\mathcal{H}$ generally includes particular insights of analysts that are exclusively available to the PM. The insights usually come as a consequence of the analysts' particular skills and efforts. Based on $\mathcal{H}$, the PM forms her (private) k-view vector $\vec{\widehat{y}}_{[k\times 1]|\mathcal{H},\mathcal{G}}$. These views are then incorporated into the return forecast vector $\vec{\widehat{m}}_{[n\times 1]} = \mathbb{E}(\vec{\tilde{r}}|\vec{\widehat{y}}, \mathcal{H}, \mathcal{G})$ with a revised covariance matrix $\widehat{\mathbf{V}}_{[n\times n]} = \mathbb{E}(\mathbf{V}|\vec{\widehat{y}}, \mathcal{H}, \mathcal{G})$. Under normality assumption, these estimates fully characterise the distribution of the security returns. Plugging $\vec{\widehat{m}}$ and $\widehat{\mathbf{V}}$ into a mean-variance optimiser, one solves a typical portfolio optimisation problem.

In practice, however, the incorporation of the public information $\mathcal{G}$ and the private information $\mathcal{H}$ in the portfolio construction process is far from trivial. Many PMs focus on exploring public market data. Various quantitative techniques have been developed. Some commonly used extrapolation techniques include historical averages, equal means, risk-adjusted equal means[6], and modern time series techniques with some prediction power.

Figure 1 shows the traditional process. Note that views are formed drawing on various information sources, and expressed in different forms, for example, explicit returns or relative performances, or even in terms of strategies. Before the BL model, systematically converting such views or information into explicit forecasts was not straightforward.

Another, more fundamental, issue is that the exploration of the commonly accessible market data might be less productive than the private information. Economists argue that these techniques would not generate insights superior to the market. In other words, if all we have are publicly available information and common techniques, then why should we not just use the market view?

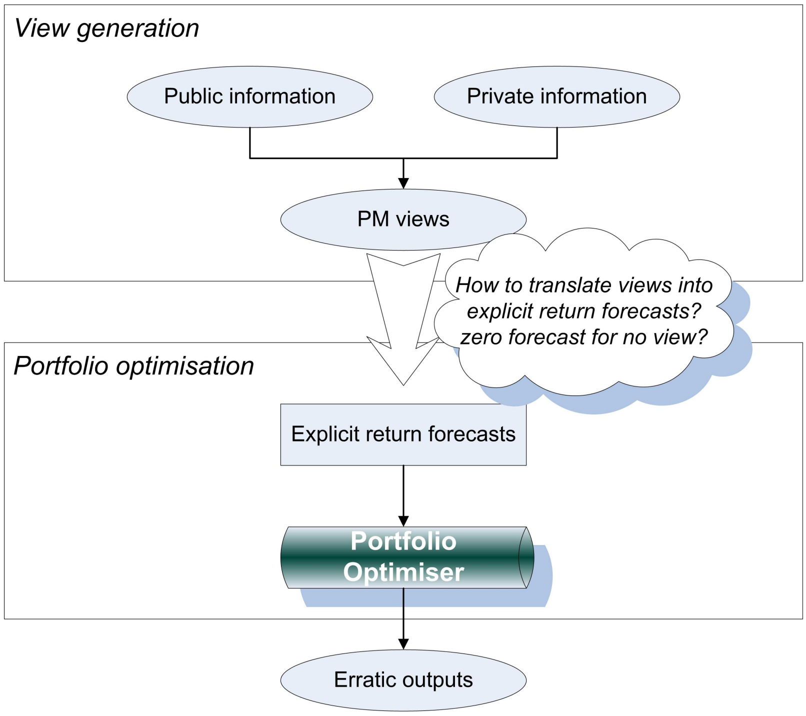
Black and Litterman (1992) take the point further and propose that without private views, the only legitimate forecasts should be backed-out from the market portfolio using the *CAPM*, the equilibrium pricer. In this case, it is optimal to simply use these forecasts to construct the portfolio and manage it passively (by holding a slice of the market portfolio). With private information, the forecasts should be updated based on the *Bayes' Rule*, the fundamental law for belief updating. Therefore, Black and Litterman recommend a Bayesian-analytic model for return forecasts and then resort to a conventional mean-variance optimiser for portfolio construction. Consequently, the model addresses the lack-of-robustness problem of portfolio optimisation through qualifying inputs rather than constraining the optimiser.

[4]In this article, $\vec{\tilde{x}}_{|\mathcal{I}}$ is used to denote our perception of $\vec{\tilde{x}}$ after examining the information $\mathcal{I}$. As in the Bayesian framework, updated perceptions are still considered random, we need such notation to distinguish them from mean (point) estimates $\mathbb{E}(\vec{\tilde{x}}|\mathcal{I})$.

[5]As is usual, we use "$\widehat{x}$" to denote an estimate for x. Note in the traditional framework, estimates are considered deterministic; whereas in the Bayesian framework, estimates are still random but with updated uncertainty.

[6]See Black and Litterman (1992) for a description and critiques of these methods.

Figure 1 The Traditional Approach to Portfolio Optimisation



3 Three Pillars: Semi-Strong Market Efficiency, CAPM, and Bayes' Rule

In the previous section, we refer to $\vec{\tilde{r}} \sim \mathbb{N}(\vec{m}, \mathbf{V})$ as the real distribution. Unlike in classical statistics in which the real means are considered deterministic (though unobservable), Bayesian statisticians consider these random themselves. Therefore, perceptions are always produced with some uncertainties, and are represented by probabilities (that is, probabilistic views). For example, after absorbing the public information $\mathcal{G}$, a PM forms her estimation $\vec{\tilde{r}}_{|\mathcal{G}} \sim \mathbb{N}(\vec{\tilde{\mu}}, \mathbf{\Sigma})$. As she is still not certain about this, particularly the mean vector $\vec{\tilde{\mu}}$, new information will always be sought for further improvements.

Note that the estimation $\vec{\tilde{r}}_{|\mathcal{G}} \sim \mathbb{N}(\vec{\tilde{\mu}}, \mathbf{\Sigma})$ only represents this investor's personal perspective. She then uses the estimation to interact with the market. In a marketplace with countless investors, there would be countless estimates at a time. Suppose they disagree on the mean vector $\vec{\tilde{\mu}}$, and so buy and sell into the market. The market 'witnesses' all these actions driven by individual predictions and prices these securities $\vec{\tilde{r}}_{M|\vec{\tilde{\mu}}_1,\vec{\tilde{\mu}}_2,\ldots} \sim \mathbb{N}(\vec{\tilde{\mu}}_M, \mathbf{\Sigma}_M)$. Unfortunately, this evaluation is beyond the ability of human beings and can only be undertaken by the market.

Thanks to the *CAPM*, one can back-out the information from the market portfolio assuming equilibrium. In other words, one may use the CAPM-assessed equilibrium estimation $\vec{\tilde{r}}_{e|\mathcal{G}} \sim \mathbb{N}(\vec{\tilde{\pi}}, \mathbf{\Sigma}_e)$ as a proxy for the market view $\vec{\tilde{r}}_{M|\vec{\tilde{\mu}}_1,\vec{\tilde{\mu}}_2,\ldots}$, where $\vec{\tilde{\pi}}$ is the vector of the CAPM-assessed excess returns[7]. Here, as the market estimation takes into account all participants' views in the whole market, one should be more confident in $\vec{\tilde{\pi}}$

[7]Unless otherwise specified, in this article, all returns refer to excess returns.

and thus the entries of $\mathbf{\Sigma}_e$ should be smaller than those of $\mathbf{\Sigma}$. It may therefore be argued that, without any superior private insights, one should prefer $\vec{\widetilde{r}}_{e|\mathcal{G}}$ to her own estimation $\vec{\widetilde{r}}_{|\mathcal{G}}$.

This basically means if all we have are public information and common techniques, the market may be smarter than us. This argument reminds us of the *semi-strong form of market efficiency hypothesis* (Fama, 1970). The assumption suggests that all public information has been absorbed into the market pricing of securities and can therefore not be explored to achieve abnormal returns. In other words, only with superior private insight and techniques can the PM make superior returns.

In this vein, the BL assessment reduces to the CAPM-equilibrium layer in the absence of any private information, where the security excess return vector can be backed-out from the market portfolio, that is, $\vec{\widetilde{\pi}}_{[n\times 1]} = \vec{\widetilde{\beta}}_{[n\times 1]}[\mathbb{E}(\widetilde{R}_M|\mathcal{G}) - r_f]$ (where $\vec{\widetilde{\beta}}$ is the vector of security exposures to the market; $\mathbb{E}(\widetilde{R}_M|\mathcal{G})$ is the expected (total) market return; and r_f is the risk-free interest rate). In the Bayesian framework, $\vec{\widetilde{\pi}}$ is considered as the *prior* estimation, since private information has not yet been examined.

After examining the private information $\mathcal{H}$, PM forms a vector of views[8], $\vec{\widetilde{y}}_{|\mathcal{H},\mathcal{G}}$[9], usually in terms of certain linear combinations of the securities in the universe, for example, strategy and comparative performance views, together with their performance estimates. These have direct implications for the securities they cover, and also, owing to the dependence structure, indirect implications for those in the universe not covered. Therefore, there is a need to update the PM's forecasts for the whole universe, and by resorting to the *Bayes' Rule*, the BL model rightly facilitates this. The updated forecasts are denoted by $\vec{\widetilde{r}}_{|\vec{\widetilde{y}},\mathcal{H},\mathcal{G}}$, and called the *posterior* estimation.

It should be noted that, in the BL model, the private information $\mathcal{H}$ is not directly fed into security return forecasts $\vec{\widetilde{r}}_{|\mathcal{H},\mathcal{G}}$; but is first absorbed into PM views $\vec{\widetilde{y}}_{|\mathcal{H},\mathcal{G}}$, and then indirectly fed into the security return forecasts. As the views are 'private', the direct security updates $\vec{\widetilde{r}}_{|\mathcal{H},\mathcal{G}}$ come only after private views, $\vec{\widetilde{y}}_{|\mathcal{H},\mathcal{G}}$, are digested by the market. Moreover, the assumed view structure as linear combination of security return forecasts is flexible enough to admit views in terms of explicit return forecasts. Therefore, the linearity assumption has broad applications.

The main contribution of the BL model is its unique insight enabling security returns to be assessed in the Bayesian framework, drawing on some views on the underlying securities or their linear combinations. Its closed-form solution for the posterior, $\vec{\widetilde{r}}_{|\vec{\widetilde{y}},\mathcal{H},\mathcal{G}}$, serves as a smooth view blender enabling PM to incorporate her view vector $\vec{\widetilde{y}}_{|\mathcal{H},\mathcal{G}}$ into the portfolio construction and rebalancing process (See Figure 2). Comparative advantages of the model include: -

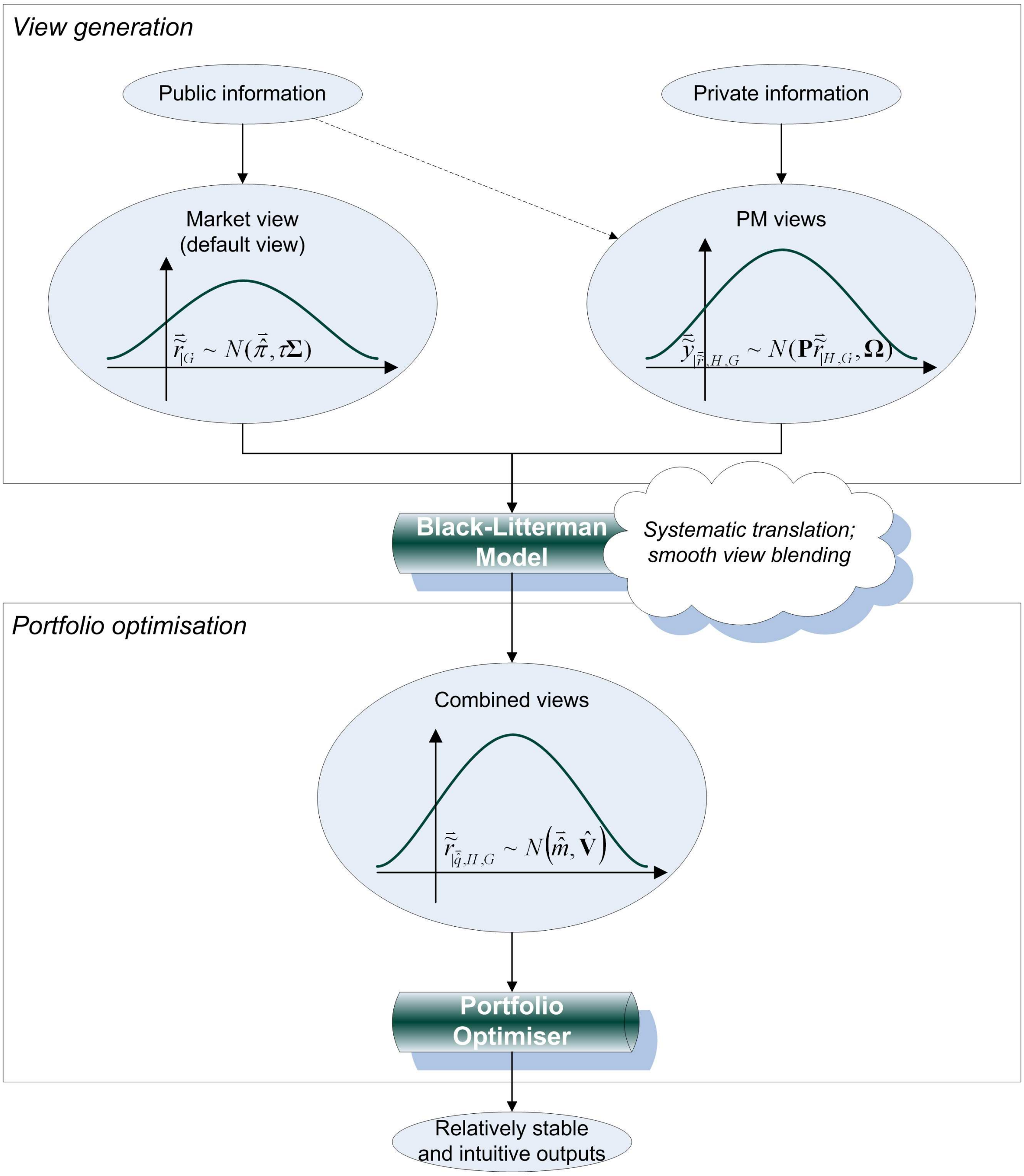
- PMs no longer need to produce forecasts for the full universe of securities. Instead, providing any k $(0 \leqslant k \leqslant n)$ number of views suffices, and these views can be *relative* (for example, Security A will outperform Security B by 2 per cent) as well as *absolute* (for example, Security A will grow by 10 per cent).

- The resultant allocation tends also to be more robust and intuitive. This is because the process starts

[8]Normally, analysts/PMs form their views according to information arrivals. In the remainder of this article, we use *view* and *information* interchangeably.

[9]In the Bayesian framework, views should also be treated as random as there are always forecast errors.

Figure 2 The BL-Based Portfolio Optimisation Framework



from a common CAPM-based *equilibrium layer*. On top of this, *tilts* are generated to reflect the private views.

Now let us delve into the formulation of the model.

4 Model Formulation and Results

The model relies on two key technical assumptions.

Assumption 4.1 (Prior Return Forecasts) *The prior return forecasts (based on the public information) are normally distributed as follows:*

$$\vec{\tilde{r}}_{|\mathcal{G}} \overset{\text{delegate}}{\rightarrow} \vec{\tilde{\mu}}_M \sim \mathbb{N}(\vec{\tilde{\pi}}, \tau\mathbf{\Sigma}) \tag{1}$$

where τ is a positive multiplier applied to the estimated covariance matrix $\mathbf{\Sigma}$ to proxy the prior error matrix.

Based on the economic reasoning in the previous section, this assumption is justified by: -

(a) With only public information $\mathcal{G}$, and supposing the market is already in equilibrium, the CAPM is the legitimate market equilibrium pricer, and thus $\vec{\tilde{\pi}}$ assesses $\vec{\tilde{r}}$.

(b) However, as the market is not necessarily in equilibrium, the assessment $\vec{\tilde{\pi}}$ suffers from errors $\mathbf{\Sigma}_e$. Using a factor model-based argument, the model assumes $\mathbf{\Sigma}_e \propto \mathbf{\Sigma}$, the estimated security covariance. Moreover, as explained in the previous section, the elements of $\mathbf{\Sigma}_e$ should be smaller than those of $\mathbf{\Sigma}$ in a market demonstrating some level of semi-strong market efficiency; thus, $\tau \leqslant 1$[10].

To assess distribution (1), we need to know the CAPM-based equilibrium return vector: $\vec{\tilde{\pi}}$. Some linear algebra drawing on the CAPM yields $\vec{\tilde{\pi}} = (\frac{\mathbb{E}(\tilde{R}_M|\mathcal{G})-r_f}{\sigma_M^2})\mathbf{\Sigma}\vec{w}_{M[n\times 1]}$. This allows us to back out security returns from the current market portfolio $\vec{w}_M$. Based on $\vec{\tilde{\pi}}$, PMs form the equilibrium layer of their portfolio and then tilt their portfolio upon arrival of new private information.

Quite realistically, PMs make investment decisions drawing on some (limited number of) theories, views or strategies. Suppose a PM has k ($\leqslant n$) private views (or theories, and so on) that are expressed or approximately represented by some linear combinations of security returns:

$$\widetilde{\mathbf{P}}_{[k\times n]} \cdot \vec{\tilde{r}} \approx \vec{\tilde{y}}_{[k\times 1]} \tag{2}$$

where $\widetilde{\mathbf{P}}$ is a matrix of view structure parameters or a vector of k strategies; and $\vec{\tilde{y}}$ represents the PM's view forecast vector.

In (2), the views/theories can somehow be qualitative, and they need to be 'calibrated' against available information.

[10]As we develop more insights in our forthcoming articles, we will see that τ is a subjective parameter and that such restriction does not exist.

Upon arrival of the private information $\mathcal{H}$, it is assumed that the PM still does not have explicit return estimates $\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}$; yet (2) materialises to the extent that $\widetilde{\mathbf{P}}_{|\mathcal{H},\mathcal{G}} \overset{\text{belief}}{\to} \mathbf{P}_{[k\times n]}$ and thus, the updated view becomes:

$$\mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}} = \vec{\tilde{y}}_{|\mathcal{H},\mathcal{G}} + \vec{\tilde{\varepsilon}}_{[k\times 1]} \tag{3}$$

where $\mathbf{P}$ is the concretised view structure matrix; $\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}$ is the (unknown but required) posterior vector of return estimates; $\vec{\tilde{y}}_{|\mathcal{H},\mathcal{G}}$ is the vector of updated view estimation; and $\vec{\tilde{\varepsilon}}$ is the vector of view estimation errors.

Note that the structure of views as linear combination of prior return forecasts is flexible enough to accommodate relative as well as absolute views regarding security returns. For example, if the only view PM has is that security A will outperform security B by 2 per cent with an uncertainty of 1 per cent (standard deviation). This relative view can be expressed as:

$$\begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \mathbb{E}(\vec{\tilde{r}}_1|\mathcal{H},\mathcal{G}) \\ \mathbb{E}(\vec{\tilde{r}}_2|\mathcal{H},\mathcal{G}) \end{pmatrix} = 2\% + \tilde{\varepsilon} \text{ where } \tilde{\varepsilon} \sim \mathbb{N}(0,(1\%)^2)$$

Alternatively, $\mathbf{P}$ can also be viewed as k strategies formulated based on the new information. In this example, 2 per cent can be considered as the expected return of the spread strategy $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

For the view estimation error vector $\vec{\tilde{\varepsilon}}$ in (3), the model makes the following normality assumption:

Assumption 4.2 (View Errors) *The view error vector is normally distributed as follows:*

$$\vec{\tilde{\varepsilon}}_{[k\times 1]} \sim \mathbb{N}(\vec{0}_{[k\times 1]}, \mathbf{\Omega}_{[k\times k]}) \tag{4}$$

where $\vec{0}$ is a vector of zeros; and $\mathbf{\Omega}$ is a diagonal variance matrix of view-estimation errors, which, for simplicity, are considered independent across views[11].

Therefore, we have $\vec{\tilde{y}}_{|\vec{\tilde{r}},\mathcal{H},\mathcal{G}} \sim \mathbb{N}(\mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}, \mathbf{\Omega})$ (conditional on $\vec{\tilde{r}}$). With our prior knowledge (1), plus these views and our final mean conviction $\vec{\tilde{y}}_{|\vec{r},\mathcal{H},\mathcal{G}} \overset{\text{belief}}{\to} \vec{\tilde{q}}_{[k\times 1]}$ (the 'ultimate' view mean estimation), the Bayes' Rule can be utilised to leverage (3) for an update of the forecasts of $\vec{\tilde{r}}$[12]. The closed-form results are as follows:

[11] This assumption comes as a convention of the classical linear regression model. This provides some simplicity. However, as far as the model is concerned, this independence restriction is not necessary.

[12] Careful readers may have noted the relationship (3) can be re-arranged and considered as a linear system $\vec{\tilde{y}}_{|\vec{\tilde{r}},\mathcal{H},\mathcal{G}} = \mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}} - \vec{\tilde{\varepsilon}}$ for fitting $\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}$, with $\vec{\tilde{q}}$ and $\mathbf{P}$ as data. It seems that $\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}$ can be estimated by the generalised least square (GLS) estimator $\left(\mathbf{P}^{\mathrm{T}}_{[n\times k]}\mathbf{\Omega}^{-1}_{[k\times k]}\mathbf{P}_{[k\times n]}\right)^{-1}\mathbf{P}^{\mathrm{T}}_{[n\times k]}\mathbf{\Omega}^{-1}_{[k\times k]}\vec{\tilde{q}}_{[k\times 1]}$. There are however two reasons why the following BL solution is more practical: (1) In practice, it is often the case that we only have k ($\leqslant n$) views. As such, the matrix $\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}$ is singular and therefore the GLS estimation is not attainable. (2) Treating $\vec{\tilde{q}}$ and $\mathbf{P}$ as data is too optimistic - these are often just analysts' opinions. The strong associated uncertainty makes the Bayesian technique a more natural fit.

Theorem 4.1 (Posterior Return Estimates) *The posterior return vector is normally distributed, that is, $\vec{\tilde{r}}_{|\vec{\hat{q}},\mathcal{H},\mathcal{G}} \sim \mathbb{N}(\vec{\widehat{m}}, \widehat{\mathbf{V}})$, where the updated mean vector is:*

$$\vec{\widehat{m}} = \left[(\tau\mathbf{\Sigma})^{-1} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}\right]^{-1}\left[(\tau\mathbf{\Sigma})^{-1}\vec{\widehat{\pi}} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\vec{\widehat{q}}\right] \tag{5}$$

and the updated variance-covariance matrix is:

$$\widehat{\mathbf{V}} = \left[(\tau\mathbf{\Sigma})^{-1} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}\right]^{-1} \tag{6}$$

Proof. See Appendix A.□

The results are intuitive. Consider the 1-security, 1-view scenario:

$$\widehat{m} = \frac{\frac{\widehat{\pi}}{\tau\sigma^2} + \frac{\widehat{q}}{\nu^2}}{\frac{1}{\tau\sigma^2} + \frac{1}{\nu^2}} = \varpi_1\widehat{\pi} + \varpi_2\widehat{q} \tag{7}$$

or

posterior estimate = confidence-weighted average of views; and

$$\frac{1}{\widehat{V}} = \frac{1}{\tau\sigma^2} + \frac{1}{\nu^2} \tag{8}$$

or

posterior confidence = aggregate of confidence in both views,

where σ is the security volatility; and ν is the view estimation error, and $\varpi_1 = \frac{\frac{1}{\tau\sigma^2}}{\frac{1}{\tau\sigma^2}+\frac{1}{\nu^2}}$ and $\varpi_2 = \frac{\frac{1}{\nu^2}}{\frac{1}{\tau\sigma^2}+\frac{1}{\nu^2}}$ are the confidence weights.

If we consider 'confidence' as the inverse of variance (the 'uncertainty'), Equation (8) basically blends the PM's confidence in the market equilibrium view, $\frac{1}{\tau\sigma^2}$, and her own view, $\frac{1}{\nu^2}$, to obtain the posterior estimation confidence as $\frac{1}{\tau\sigma^2} + \frac{1}{\nu^2}$. This simple additive relationship comes as a result of the Bayes's Rule. The more confident the PM is in either view, the more confident she will be in her posterior (8).

Equation (7) simply combines the market equilibrium with the PM views through a confidence-weighted average scheme. Again, this relationship is dictated by the Bayes' Rule. It is easy to see that the more confident the PM is about her view $\widehat{q}$, the more weight she should put on the view, and therefore her posterior forecasts should be adjusted more towards $\widehat{q}$. Consequently, the more she should tilt her new portfolio towards reflecting the view. Contrarily, the less confident, the more she should rely on the market portfolio. In the extreme case in which her confidence in the views is minimal (or she has no views), Theorem 4.1 reduces to (1) and she should do nothing but hold the prior portfolio.

5 Implementation

The BL framework in practice involves four steps:

Step 1: *Data collection for the market-wide variables, portfolio-specific variables, and the PM views.*

The following inputs are required:

General economy-related input:

r_f: the risk-free interest rate

Benchmark portfolio-related inputs: pick a prior portfolio, with which the universe and the following are obtained:

$\vec{w}_M$: the market portfolio weight vector

$\mathbb{E}(\widetilde{R}_M|\mathcal{G})$: the estimated gross market portfolio return

Portfolio-related inputs:

τ: the risk multiplier. The value choice for this parameter should be considered together with the view-uncertainty matrix $\mathbf{\Omega}$ to achieve a desired balance for view shinkage

$\mathbf{\Sigma}$: the estimated security return variance-covariance matrix

View-related inputs:

$\mathbf{P}$: the view strategy structure

$\vec{\widetilde{q}}$: the view estimates, that is, estimated strategy returns

$\mathbf{\Omega}$: the diagonal matrix containing the view variances

Step 2: *Back-out the market view from the market portfolio based on the CAPM.*

This basically requires the assessment of the CAPM-based excess returns $\vec{\widetilde{\pi}} = \frac{\mathbb{E}(\widetilde{R}_M|\mathcal{G})-r_f}{\vec{w}_M^{\mathrm{T}}\mathbf{\Sigma}\vec{w}_M}\mathbf{\Sigma}\vec{w}_M$.

Step 3: *Update the forecasts according to Theorem 4.1.*

This involves substituting the results from Step 2, together with other inputs, into Theorem 4.1 to evaluate the posterior mean $\vec{\widetilde{m}}$ and error matrix $\widehat{\mathbf{V}}$.

Step 4: *Optimise the allocation based on the posterior estimates to decide how to tilt from the market portfolio.*

The general mean-variance optimisation problem is:

$$\max_{\vec{w}}\{\vec{w}^{\mathrm{T}}\vec{\widetilde{m}} - \lambda\vec{w}^{\mathrm{T}}\widehat{\mathbf{V}}\vec{w}\} \tag{9}$$

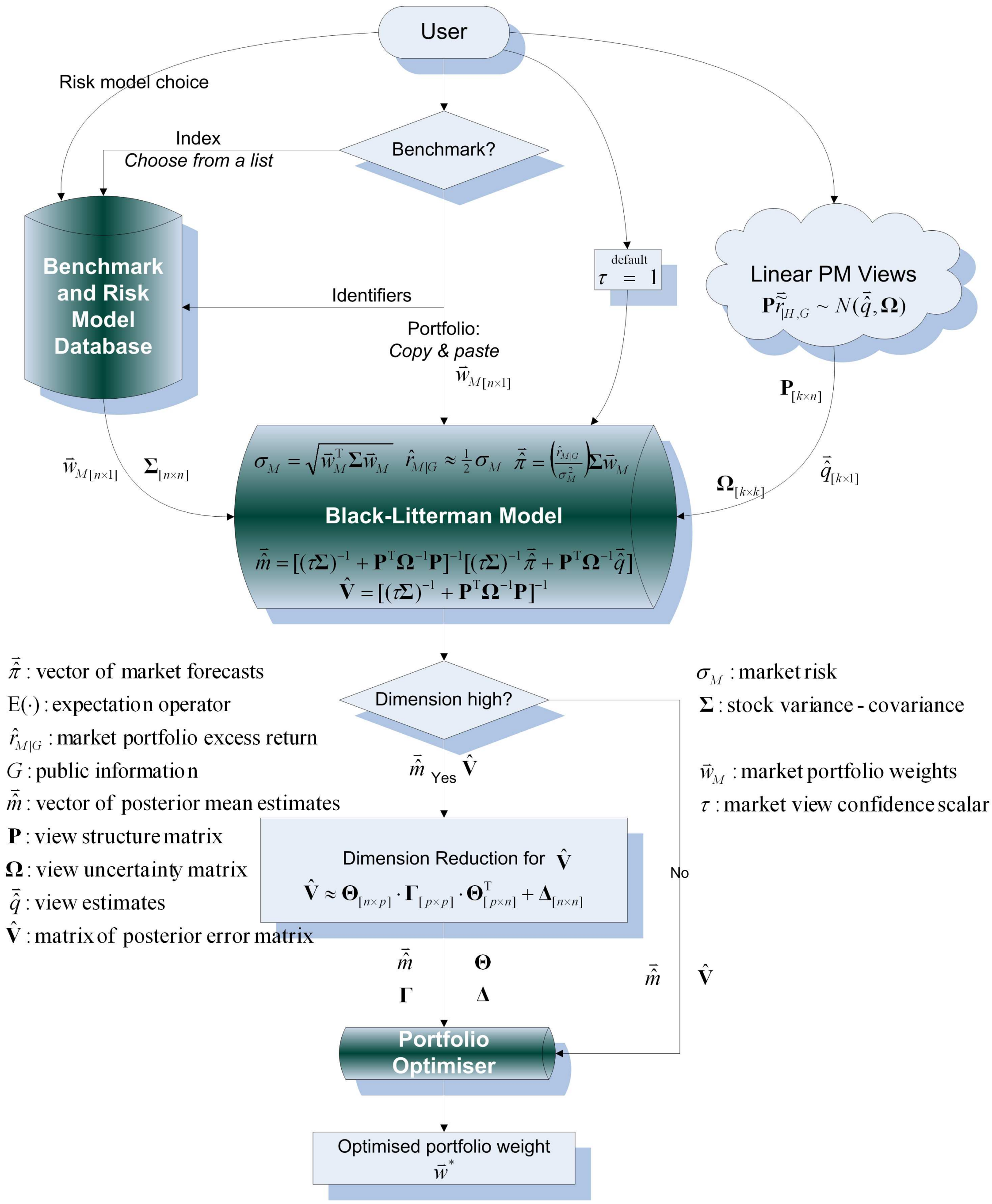
Since most of the data items can be obtained from a market database with a risk model, the PM only needs to concentrate on the view-related inputs: τ, $\mathbf{P}$, $\vec{\widetilde{q}}$ and $\mathbf{\Omega}$.

Figure 3 illustrates how the user should prepare inputs and how the model interacts with other components in a typical equity analytical framework.

6 Practical Issue: Curse of Dimensionality

Note that the posterior covariance matrix $\widehat{\mathbf{V}}_{[n\times n]}$ produced by the BL model is a fully populated numerical matrix. Recall that the BL model was originally proposed for low-dimension global asset allocation. This

Figure 3 Interactions between Different Components in the BL Framework



application avoids the dimensionality issue. However, to an equity portfolio manager who faces a large universe with say, $n > 1000$, this 'heavy' matrix poses significant computational challenge to any quadratic optimiser. In order to leverage the capacity of the optimiser, dimension reduction will be needed in practice.

We recommend the following eigensystem analysis. Since $\widehat{\mathbf{V}}$ is real, symmetric and positive-definite, the following diagonalisation is guaranteed:

$$\widehat{\mathbf{V}} = \mathbf{E}_{[n\times n]}\mathbf{D}_{[n\times n]}\mathbf{E}^{\mathrm{T}}_{[n\times n]} \tag{10}$$

where $\mathbf{D}$ is the diagonal matrix with the eigenvalues sorted from high to low; and $\mathbf{E}$ represents the loadings matrix constructed by the eigenvectors corresponding to the sorted eigenvalues.

Using the first p ($p \ll n$) eigen components to approximate the original matrix and collect residuals, we have:

$$\widehat{\mathbf{V}} = \mathbf{\Theta}_{[n\times p]}\mathbf{\Gamma}_{[p\times p]}\mathbf{\Theta}^{\mathrm{T}}_{[p\times n]} + \mathbf{R}_{[n\times n]} \tag{11}$$

where $\mathbf{\Gamma}$ is the diagonal matrix reduced from $\mathbf{D}$ by retaining the first p engenvalues; $\mathbf{\Theta}$ is the loading matrix reduced from $\mathbf{E}$ by retaining the first p eigenvectors; and $\mathbf{R}$ represents the residual matrix.

Impose the following treatment:

$$\widehat{\mathbf{V}} \approx \mathbf{\Theta}\mathbf{\Gamma}\mathbf{\Theta}^{\mathrm{T}} + \mathbf{\Delta}_{[n\times n]} \tag{12}$$

where $\mathbf{\Delta}$ is the diagonal matrix obtained from $\mathbf{R}$ by setting all its off-diagonal elements to 0.

By such approximation, we attain a sparse representation of $\widehat{\mathbf{V}}$ (that is, $\mathbf{\Theta}$ is just a $[n \times p]$ matrix, $\mathbf{\Gamma}$ is a small $[p \times p]$ matrix, and $\mathbf{\Delta}$ is a diagonal matrix). This will significantly help the optimiser.

With the same choice of p, the approximation has exactly the same quality as a principal component analysis (PCA) risk model. In risk modelling, the application of the eigensystem analysis is meant for noise disposal. So generally, a very limited number l of principal components enter in $\mathbf{\Theta}\mathbf{\Gamma}\mathbf{\Theta}^{\mathrm{T}}$. In our case, however, the main purpose is to reduce the dimension to the level that the optimiser can handle. Therefore, the choice of p is only subject to the optimiser's capacity, which can easily handle several hundreds. We may therefore set p to a number much larger than l (for example, $p = 100$) to allow a very good approximation of the posterior covariance matrix.

Alternatively, to resolve the dimension issue, one may also rank the securities according to the posterior estimates and reduce the universe before optimisation.

7 Conclusion

This article is an introduction to the BL Model. Essential technical details, that is, motivation, intellectual roots, model formulation, and implementation guidance, are included with a view to enhancing appreciation and enabling implementation of the model.

In addressing the lack-of-robustness issue of the mean-variance optimiser, the BL model aims to qualify the view inputs rather than to constrain the optimiser. A market that demonstrates *semi-strong market* *efficiency* is smarter than any individual with only public information and conventional technique; therefore, one should simply rely on the market view, implied from the market portfolio by the *CAPM*, the equilibrium pricer. Only when an investor has unique insight and superior forecasting technique can she potentially outperform the market. To take advantage of these, the *Bayes' Rule*, the fundamental Law of belief updating, is employed for view processing (that is, updating, translation, blending and shrinkage). It is based on these three notions that the BL model is established.

Drawing on normality and linearity assumptions, this model admits analytical solutions. Implementation can be efficient. As the outputs are in the form of explicit return forecasts together with a covariance matrix, a standard mean-variance optimiser can produce allocation recommendation which is considered robust owing to the shrinkage effect. This emancipates PMs from the job of view processing and portfolio construction, enabling them to concentrate on alpha generation.

However, the user should be aware of some BL practical issues.

- ✓ *Curse of dimensionality*. The BL model spits out a fully populated numerical posterior matrix. In case of a large universe, this places significant burden on the optimiser. We resolve this issue by obtaining a sparse representation of the matrix drawing on the principal component approximation.
- ☐ *Confidence parameter setting*. Setting τ and $\mathbf{\Omega}$ is a classical issue inherited from the Bayesian framework.
- ☐ *View correlations*. For simplicity, the BL model treats views as orthogonal. In reality, view correlations may exist but are hard to quantify. Is this a problem?
- ☐ *Prior setting*. In practice, there are a variety of investment styles. Equilibrium is an abstract concept, and the market portfolio does not seem to be a suitable starting point for all strategies. How should we choose an appropriate prior?
- ☐ *Factor-based portfolio construction*. How can we apply the BL technique to the popular Fama-French factor ranking approach?
- ☐ *Optimiser issues*. The BL framework calls an optimiser for allocation purpose. Whereas the optimiser behaves as a 'black box'. There are risk-aversion as well as transparency issues. Can we make the framework more open and intuitive?
- ☐ *Risk model quality*. The risk model itself has errors. How can we economically use it, or make reference to more than one risk models in allocation?
- ☐ *Linearity and normality assumptions*. In the BL model, security returns are considered normal; and the factor model linear. How to apply this model to non-normal, non-linear markets?

In our forthcoming articles, we aim to make the BL technique practical by resolving all these issues.

References

[1] Black, F. and R. Litterman (1992), "Global Portfolio Optimisation" in *Financial Analysts Journal*, Sep./Oct. 1992, Vol.48, No.5, pp. 28-43

[2] Fama, E. F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work" in *Journal of Finance*, Vol.25, No.2, pp.383-417

[3] Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press: Princeton, NJ

[4] He, G. and R. Litterman (2003), "The Intuition Behind Black-Litterman Model Portfolios" Available at SSRN: http://ssrn.com/abstract=334304 or DOI: 10.2139/ssrn.334304

[5] Idzorek, T. M. (2004), "A Step-by-Step Guide to the Black-Litterman Model", Zephyr Working Document 2004

[6] Jones, R., T. Lim and P. J. Zangari (2007), "The Black-Litterman Model for Structured Equity Portfolios" in *Journal of Portfolio Management*, Winter 2007, pp.24-33

[7] Lee, W. (2000), *Advanced Theory and Methodology of Tactic Asset Allocation*, forth-coming book

[8] Martellini L. and V. Ziemann (2007) "Extending Black-Litterman Analysis Beyond the Mean-Variance Framework", Working Document, EDHEC

[9] Meucci, A. (2005), "Beyond Black-Litterman: Views on Non-normal Market". Available at SSRN: http://ssrn.com/abstract=848407

[10] Satchell, S. and A. Scowcroft (2000), "A Demystification of the Black-Litterman Model: Managing Quantitative and Traditional Portfolio Construction", in *Journal of Asset Management*, Vol.1, No.2, pp.138-50

[11] Zhou, G. (2008), "An Extension of the Black-Litterman Model: Letting the Data Speak", Working Document, Washington University in St. Louis

Appendix A: Proof

Theorem 4.1 (Posterior Return Estimates) *The posterior return vector is normally distributed, that is,* $\vec{\tilde{r}}_{|\vec{\tilde{q}},\mathcal{H},\mathcal{G}} \sim \mathbb{N}(\vec{\tilde{m}}, \widehat{\mathbf{V}})$*, where the updated mean vector is:*

$$\vec{\tilde{m}} = \left[(\tau\mathbf{\Sigma})^{-1} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}\right]^{-1}\left[(\tau\mathbf{\Sigma})^{-1}\vec{\tilde{\pi}} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\vec{\tilde{q}}\right] \tag{13}$$

and the updated variance-covariance matrix is:

$$\widehat{\mathbf{V}} = \left[(\tau\mathbf{\Sigma})^{-1} + \mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}\right]^{-1} \tag{14}$$

Proof. We first deal with a general case and then apply it into the BL context. In general, assume the following linear relationship:

$$\vec{\tilde{Y}}_{[k\times 1]} = \mathbf{Q}_{[k\times n]}\vec{\tilde{X}}_{[n\times 1]} - \vec{\tilde{u}}_{[k\times 1]}, \; (k \leqslant n) \tag{15}$$

where $\mathbf{Q}$ is a known/observable matrix; $\vec{\tilde{Y}}$ is an observable vector; $\vec{\tilde{X}}$ is unobservable and needs to be estimated; and $\vec{\tilde{u}}$ is the error vector.

Also, assume the following Gaussian distributions for regression errors and the prior:

$$\vec{\tilde{u}} \sim \mathbb{N}(\vec{0}_{[k\times 1]}, \mathbf{\Lambda}_{[k\times k]}); \text{ and} \tag{16}$$

$$\vec{\tilde{X}} \sim \mathbb{N}(\vec{\mu}_{X[n\times 1]}, \mathbf{\Sigma}_{X[n\times n]}) \tag{17}$$

We therefore have, conditional on a realisation of $\vec{\tilde{X}}$,

$$\vec{\tilde{Y}}_{|\vec{\tilde{X}}} = \mathbf{Q}\vec{\tilde{X}} - \vec{\tilde{u}} \sim \mathbb{N}(\mathbf{Q}\vec{\tilde{X}}, \mathbf{\Lambda}) \tag{18}$$

In terms of *probability density function* (*pdf*), distributions (17) and (18) can be equivalently written as:

$$f(\vec{\tilde{X}}) = \frac{1}{(2\pi)^{\frac{n}{2}}}|\mathbf{\Sigma}_X|^{-\frac{1}{2}}e^{-\frac{1}{2}(\vec{\tilde{X}}-\vec{\mu}_X)^{\mathrm{T}}(\mathbf{\Sigma}_X)^{-1}(\vec{\tilde{X}}-\vec{\mu}_X)} \tag{19}$$

where $|\cdot|$ gives the determinant of the matrix it applies to; and

$$f(\vec{\tilde{Y}}|\vec{\tilde{X}}) = \frac{1}{(2\pi)^{\frac{k}{2}}}|\mathbf{\Lambda}|^{-\frac{1}{2}}e^{-\frac{1}{2}(\vec{\tilde{Y}}-\mathbf{Q}\vec{\tilde{X}})^{\mathrm{T}}(\mathbf{\Lambda})^{-1}(\vec{\tilde{Y}}-\mathbf{Q}\vec{\tilde{X}})}, \tag{20}$$

respectively.

We try to imply the probability distribution of $\vec{\tilde{X}}_{|\vec{\tilde{Y}}}$ from the joint *pdf*:

$$\begin{aligned} f(\vec{\tilde{X}}, \vec{\tilde{Y}}) &= f(\vec{\tilde{X}})f(\vec{\tilde{Y}}|\vec{\tilde{X}}) \\ &= \frac{1}{(2\pi)^{\frac{n+k}{2}}}|\mathbf{V}_a|^{-\frac{1}{2}}e^{-\frac{1}{2}\vec{\alpha}^{\mathrm{T}}\mathbf{V}_a^{-1}\vec{\alpha}} \end{aligned} \tag{21}$$

where

$$\vec{\alpha}_{[(n+k)\times 1]} = \begin{pmatrix} \vec{\tilde{X}} - \vec{\mu}_X \\ \vec{\tilde{Y}} - \mathbf{Q}\vec{\tilde{X}} \end{pmatrix} \tag{22}$$

is considered as the *error vector*; and

$$\mathbf{V}_{a[(n+k)\times(n+k)]} = \begin{pmatrix} \mathbf{\Sigma}_X & \mathbf{0}_{[n\times k]} \\ \mathbf{0}_{[k\times n]} & \mathbf{\Lambda}_{[k\times k]} \end{pmatrix} \tag{23}$$

We hope, based on the Bayes' Rule, it is possible to express (21) in terms of $f(\vec{\tilde{X}}|\vec{\tilde{Y}})f(\vec{\tilde{Y}})$ since we are interested in the posterior estimation of $\vec{\tilde{X}}$ given $\vec{\tilde{Y}}$. This translates into a need to replace the dependence of the mean estimates of $\vec{\tilde{Y}}$ on $\vec{\tilde{X}}$ (as in (18)) with a dependence of the mean estimates of $\vec{\tilde{X}}$ on $\vec{\tilde{Y}}$. In other words, through some transformation, we need to get rid of $\vec{\tilde{X}}$ from the lower part of the error vector in (22), but allow $\vec{\tilde{Y}}$ to enter the upper part. To this end, we construct the following matrix (Hamilton, 1994, Ch.12):

$$\begin{aligned} & \mathbf{A}_{[(n+k)\times(n+k)]} \\ = & \begin{pmatrix} \mathbf{I}_{[n\times n]} & -[(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1} \\ \mathbf{0}_{[k\times n]} & \mathbf{I}_{[k\times k]} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{[n\times n]} & \mathbf{0}_{[n\times k]} \\ \mathbf{Q}_{[k\times n]} & \mathbf{I}_{[k\times k]} \end{pmatrix} \\ = & \begin{pmatrix} [(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}(\mathbf{\Sigma}_X)^{-1} & -[(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}\mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1} \\ \mathbf{Q}_{[k\times n]} & \mathbf{I}_{[k\times k]} \end{pmatrix} \end{aligned} \tag{24}$$

where note $|\mathbf{A}| = 1$.

Using $\mathbf{A}$ as the transform matrix, we define:

$$\begin{aligned} \vec{\alpha}' &= \mathbf{A}\vec{\alpha} \\ &= \begin{pmatrix} \vec{\tilde{X}} - [(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}[(\mathbf{\Sigma}_X)^{-1}\vec{\mu}_X + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\vec{\tilde{Y}}] \\ \vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X \end{pmatrix} \\ &= \begin{pmatrix} \vec{\tilde{X}} - \vec{\tilde{m}}(\mathbf{Q}, \mathbf{\Sigma}_X, \mathbf{\Lambda}, \vec{\mu}_X, \vec{\tilde{Y}}) \\ \vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X \end{pmatrix} \end{aligned} \tag{25}$$

where

$$\vec{\tilde{m}}(\mathbf{Q}, \mathbf{\Sigma}_X, \mathbf{\Lambda}, \vec{\mu}_X, \vec{\tilde{Y}}) = [(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}[(\mathbf{\Sigma}_X)^{-1}\vec{\mu}_X + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\vec{\tilde{Y}}];$$

and

$$\begin{aligned} \mathbf{V}_a' &= \mathbf{A}\mathbf{V}_a\mathbf{A}^{\mathrm{T}} \\ &= \begin{pmatrix} [(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1} & \mathbf{0}_{[n\times k]} \\ \mathbf{0}_{[k\times n]} & \mathbf{Q}\mathbf{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \mathbf{\Lambda} \end{pmatrix} \\ &= \begin{pmatrix} \widehat{\mathbf{V}}(\mathbf{Q}, \mathbf{\Sigma}_X, \mathbf{\Lambda}) & \mathbf{0}_{[n\times k]} \\ \mathbf{0}_{[k\times n]} & \mathbf{Q}\mathbf{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \mathbf{\Lambda} \end{pmatrix} \end{aligned} \tag{26}$$

where $\widehat{\mathbf{V}}(\mathbf{Q}, \mathbf{\Sigma}_X, \mathbf{\Lambda}) = [(\mathbf{\Sigma}_X)^{-1} + \mathbf{Q}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Q}]^{-1}$.

Therefore, (21) can be rearranged, noting $|\mathbf{A}^{-1}| = |\mathbf{A}| = 1$, as follows:

$$\begin{aligned}
& f(\vec{\tilde{X}}, \vec{\tilde{Y}}) \\
= \; & \frac{1}{(2\pi)^{\frac{n+k}{2}}} |\mathbf{V}_a|^{-\frac{1}{2}} e^{-\frac{1}{2}\vec{\alpha}^{\mathrm{T}}\mathbf{V}_a^{-1}\vec{\alpha}} \\
= \; & \frac{1}{(2\pi)^{\frac{n+k}{2}}} |(\mathbf{A}^{-1})\mathbf{V}'_a(\mathbf{A}^{-1})^{\mathrm{T}}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{A}^{-1}\vec{\alpha}')^{\mathrm{T}}[(\mathbf{A}^{-1})\mathbf{V}'_a(\mathbf{A}^{-1})^{\mathrm{T}}]^{-1}(\mathbf{A}^{-1}\vec{\alpha}')} \\
= \; & \frac{1}{(2\pi)^{\frac{n+k}{2}}} |\mathbf{V}'_a|^{-\frac{1}{2}} e^{-\frac{1}{2}(\vec{\alpha}')^{\mathrm{T}}(\mathbf{V}'_a)^{-1}(\vec{\alpha}')} \qquad (27)
\end{aligned}$$

$$\begin{aligned}
\therefore \quad & f(\vec{\tilde{X}}, \vec{\tilde{Y}}) \\
= \; & \frac{1}{(2\pi)^{\frac{n+k}{2}}} \begin{vmatrix} \widehat{\mathbf{V}}(\cdot) & \mathbf{0}_{[n\times k]} \\ \mathbf{0}_{[k\times n]} & \mathbf{Q}\boldsymbol{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \boldsymbol{\Lambda} \end{vmatrix}^{-\frac{1}{2}} \cdot \\
& e^{-\frac{1}{2}\begin{pmatrix} \vec{\tilde{X}} - \vec{\tilde{m}}(\cdot) \\ \vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \widehat{\mathbf{V}}^{-1}(\cdot) & \mathbf{0}_{[n\times k]} \\ \mathbf{0}_{[k\times n]} & (\mathbf{Q}\boldsymbol{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \boldsymbol{\Lambda})^{-1} \end{pmatrix} \begin{pmatrix} \vec{\tilde{X}} - \vec{\tilde{m}}(\cdot) \\ \vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X \end{pmatrix}} \\
= \; & \frac{1}{(2\pi)^{\frac{n}{2}}} \left|\widehat{\mathbf{V}}(\cdot)\right|^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\vec{\tilde{X}} - \vec{\tilde{m}}(\cdot)\right)^{\mathrm{T}} \widehat{\mathbf{V}}^{-1}(\cdot)\left(\vec{\tilde{X}} - \vec{\tilde{m}}(\cdot)\right)} \cdot \\
& \frac{1}{(2\pi)^{\frac{k}{2}}} \left|\mathbf{Q}\boldsymbol{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \boldsymbol{\Lambda}\right|^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X\right)^{\mathrm{T}} \left(\mathbf{Q}\boldsymbol{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \boldsymbol{\Lambda}\right)^{-1}\left(\vec{\tilde{Y}} - \mathbf{Q}\vec{\mu}_X\right)} \qquad (28)
\end{aligned}$$

where $\widehat{\mathbf{V}}(\cdot) = \widehat{\mathbf{V}}(\mathbf{Q}, \boldsymbol{\Sigma}_X, \boldsymbol{\Lambda})$; and $\vec{\tilde{m}}(\cdot) = \vec{\tilde{m}}(\mathbf{Q}, \boldsymbol{\Sigma}_X, \boldsymbol{\Lambda}, \vec{\mu}_X, \vec{\tilde{Y}})$.

From (15)-(17), we have the unconditional distribution:

$$\vec{\tilde{Y}} \sim \mathbb{N}\left(\mathbf{Q}\vec{\mu}_X, \mathbf{Q}\boldsymbol{\Sigma}_X\mathbf{Q}^{\mathrm{T}} + \boldsymbol{\Lambda}\right) \qquad (29)$$

With (29), it is easy to imply from (28) the following: -

$$\vec{\tilde{X}}_{|\vec{\tilde{Y}}} \sim \mathbb{N}\left(\vec{\tilde{m}}(\mathbf{Q}, \boldsymbol{\Sigma}_X, \boldsymbol{\Lambda}, \vec{\mu}_X, \vec{\tilde{Y}}), \widehat{\mathbf{V}}(\mathbf{Q}, \boldsymbol{\Sigma}_X, \boldsymbol{\Lambda})\right) \qquad (30)$$

To assess $\vec{\tilde{X}}_{|\vec{\tilde{Y}}}$ in (30), we use our best knowledge regarding $\mathbf{Q}$, $\boldsymbol{\Sigma}_X$, $\boldsymbol{\Lambda}$, $\vec{\mu}_X$, and $\vec{\tilde{Y}}$.

Recall in the model setting, our best knowledge based on the public information $\mathcal{G}$ leads to the following prior belief:

$$\vec{\tilde{r}}_{|\mathcal{G}} \sim \mathbb{N}(\vec{\tilde{\pi}}, \tau\boldsymbol{\Sigma}) \qquad (31)$$

After examining the private information $\mathcal{H}$, we form the following updated views:

$$\mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}} = \vec{\tilde{y}}_{|\mathcal{H},\mathcal{G}} + \vec{\tilde{\varepsilon}} \qquad (32)$$

where $\mathbf{P}$ is the view structure; $\vec{\tilde{y}}_{|\mathcal{H},\mathcal{G}}$ is the view forecast vector; and we assume $\vec{\tilde{\varepsilon}} \sim \mathbb{N}(\vec{0}, \boldsymbol{\Omega})$.

Therefore, conditional on a realisation of $\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}$:

$$\vec{\tilde{y}}_{|\vec{r},\mathcal{H},\mathcal{G}} = \mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}} - \vec{\tilde{\varepsilon}} \sim \mathbb{N}(\mathbf{P}\vec{\tilde{r}}_{|\mathcal{H},\mathcal{G}}, \boldsymbol{\Omega}) \qquad (33)$$

Substituting $\vec{\tilde{r}}$ for $\vec{\tilde{X}}$ and $\vec{\tilde{y}}$ for $\vec{\tilde{Y}}$ into (30), we reach:

$$\begin{aligned}
& \vec{\tilde{r}}_{|\vec{\tilde{y}},\mathcal{H},\mathcal{G}} \\
\sim \quad & \mathbb{N}\left(\vec{\tilde{m}}(\mathbf{P},\tau\mathbf{\Sigma},\mathbf{\Omega},\vec{\tilde{\pi}},\vec{\tilde{y}}_{|\vec{\tilde{r}},\mathcal{H},\mathcal{G}}),\widehat{\mathbf{V}}(\mathbf{P},\tau\mathbf{\Sigma},\mathbf{\Omega})\right) \\
= \quad & \mathbb{N}\left([(\tau\mathbf{\Sigma})^{-1}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}]^{-1}[(\tau\mathbf{\Sigma})^{-1}\vec{\tilde{\pi}}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\vec{\tilde{y}}_{|\vec{\tilde{r}},\mathcal{H},\mathcal{G}}],[(\tau\mathbf{\Sigma})^{-1}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}]^{-1}\right) \qquad (34)
\end{aligned}$$

Finally, using our eventual conviction about the mean of $\vec{\tilde{y}}_{|\vec{\tilde{r}},\mathcal{H},\mathcal{G}} \stackrel{\text{belief}}{\to} \vec{\tilde{q}}$, we reach the following posterior belief:

$$\begin{aligned}
& \vec{\tilde{r}}_{|\vec{\tilde{q}},\mathcal{H},\mathcal{G}} \\
\sim \quad & \mathbb{N}\left([(\tau\mathbf{\Sigma})^{-1}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}]^{-1}[(\tau\mathbf{\Sigma})^{-1}\vec{\tilde{\pi}}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\vec{\tilde{q}}],[(\tau\mathbf{\Sigma})^{-1}+\mathbf{P}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{P}]^{-1}\right) \qquad (35)
\end{aligned}$$

This completes the proof. $\square$