The Black-Litterman Model: A Detailed Exploration
Revised: May 2, 2008
© Copyright 2008 - Jay Walters, CFA
jwalters@computer.org

Abstract

This paper provides a description of the Black-Litterman model. It reviews the model as defined by the original authors, and then discusses the results found in several of the widely accessible papers on the subject. Several extensions to the model are considered, and where possible examples worked illustrating the concepts embodied in the extensions.

Overview

The paper begins with an overview of the Black-Litterman model. Next it proceeds to work through the various pieces of the model, the reference model for returns, the prior distribution and the investors views. Theil's mixed estimation model is described and then the mixing model is also viewed from the Bayesian perspective. A reference implementation of Black-Litterman is discussed, followed by a section comparing the reference implementation to several author's results. Some extensions to the Black Litterman model are explained and Idzorek's example is fully worked. A brief annotated bibliography is provided to complete the paper.

Black-Litterman the Model

The Black-Litterman model was first published by Fischer Black and Robert Litterman of Goldman Sachs in an internal Goldman Sachs Fixed Income document in 1990. This paper was then published in the Journal of Fixed Income in 1991. A longer and richer paper was published in 1992 in the Financial Analysts Journal (FAJ). The latter article was then republished by FAJ in the mid 1990's. Copies of the FAJ article are widely available on the Internet. It provides the rationale for the methodology, and some information on the derivation, but does not show all the formulas or a full derivation. It also includes a rather complex worked example, which is difficult to reproduce due to the number of assets and use of currency hedging.

The Black-Litterman model makes two significant contributions to the problem of asset allocation. First, it provides an intuitive prior, the CAPM equilibrium market portfolio, as a starting point for estimation of asset returns. Previous similar work started either with the uninformative uniform prior distribution or with the global minimum variance portfolio. The latter method, described by Frost and Savarino (1986) and Jorion (1986), took a shrinkage approach to improve the final asset allocation. Neither of these methods has an intuitive connection back to the market,. The idea that one could use 'reverse optimization' to generate a stable distribution of returns from the CAPM market portfolio as a starting point is a significant improvement to the process of return estimation.

Second, the Black-Litterman model provides a clear way to specify investors views and to blend the investors views with prior information. The investor's views are allowed to be partial or complete, and the views can span arbitrary and overlapping sets of assets. The model estimates expected excess

1 The author gratefully acknowledges feedback and comments from Attilio Meucci and Boris Gnedenko,
returns and covariances which can be used as input to an optimizer. Prior to their paper, nothing similar had been published. The mixing process had been studied, but nobody had applied it to the problem of estimating returns. No research linked the process of specifying views to the blending of the prior and the investors views. The Black-Litterman model provides a quantitative framework for specifying the investor's views, and a clear way to combine those investor's views with an intuitive prior to arrive at a new combined distribution.

When used as part of an asset allocation process, the Black-Litterman model leads to more stable and more diversified portfolios than plain mean-variance optimization. Using this model requires a large amount of data, some of which may be hard to find. First, the investor needs to identify their investable universe and find the market capitalization of each asset class. Then, they need to identify a time series of returns for each asset class, and for the risk free asset in order to compute a covariance matrix of excess returns. Often a proxy will be used for the asset class, such as using a representative index, e.g. S&P 500 Index for US Domestic large cap equities. The return on a short term sovereign bond, e.g US 13-week treasury bill, would suffice for most investor's risk free rate. Finally, the investor needs to quantify their views so that they can be applied and new return estimates computed. Finding the market capitalization information for liquid asset classes might be a challenge for an individual investor, but likely presents little obstacle for an institutional investor because of their access to index information from the various providers. Given the limited availability of market capitalization data for illiquid asset classes, e.g. real estate, private equity, even institutional investors might have a difficult time piecing together adequate market capitalization information. Finally, the outputs from the model need to be fed into a portfolio optimizer to generate the efficient frontier, and an efficient portfolio selected. Bevan and Winkelmann (1999) provide a description of their asset allocation process (for international fixed income) and how they use the Black-Litterman model within that process. This includes their approaches to calibrating the model and information on how they compute the Covariance matrices.

The standard Black-Litterman model does not provide direct sensitivity to market factors besides the asset returns, however Krishnan and Mains (2005) have provided extensions to the model which allow adding additional cross asset class factors such as a recession, or credit, market factor.

Most of the Black-Litterman literature reports results using the closed form solution for unconstrained optimization. They also tend to use non-extreme views in their examples. I believe this is done for simplicity, but it is also a testament to the stability of the outputs of the Black-Litterman model that useful results can be generated via this process. As part of a investment process, it is reasonable to conclude that some constraints would be applied at least in terms of restricting short selling and limiting concentration in asset classes. Lack of a budget constraint is also consistent with a Bayesian investor who may not wish to be 100% invested in the market due to uncertainty about their beliefs in the market.

For the ensuing discussion, we will describe the CAPM equilibrium distribution as the prior distribution, and the investors views as the conditional distribution. This is consistent with the original Black and Litterman (1992) paper. It also is consistent with our intuition about the outcome in the absence of a conditional distribution (no views in Black-Litterman terminology.) This is the opposite of the way most examples of Bayes Theorem are defined, they start with a non-statistical prior distribution, and then add a sampled (statistical) distribution of new data as the conditional distribution. The mixing model we will use, and our use of normal distributions, will bring us to the same outcome independent of these choices.
The Reference Model

The reference model for returns is the base upon which the rest of Black-Litterman is built. It includes the assumptions about which variables are random, and which are not. It also defines which parameters are modeled, and which are not modeled. Most importantly, the reference model is not obvious and it is easy to miss the important details.

We start with normally distributed expected returns

\[ E(r) \sim N(\mu, \Sigma) \]

The fundamental goal of the Black-Litterman model is to model these expected returns, which are assumed to be normally distributed with mean \( \mu \) and variance \( \Sigma \). Note that we will need both of these values, the expected returns and covariance matrix later as inputs into a Mean-Variance optimization.

Black-Litterman defines \( \mu \), the mean return, as a random variable itself distributed as

\[ \mu \sim N(\pi, \Sigma_\pi) \]

\( \pi \) is our estimate of the mean and \( \Sigma_\pi \) is the variance of our estimate from the mean return \( \mu \). Another way to view this simple linear relationship is shown in the formula below.

(1) \[ \pi = \mu + \epsilon \]

\( \epsilon \) is normally distributed with mean 0 and variance \( \Sigma_\epsilon \). \( \Sigma_\epsilon \) is assumed to be uncorrelated with \( \Sigma \), the variance of returns about \( \mu \).

Formula (1) may seem to be incorrect with \( \pi \) on the left hand side, however \( \pi \) varies around \( \mu \) so it is correct.

We can complete the reference model by defining \( \Sigma_r \) as the variance of our estimate \( \pi \). From formula (1) and the assumption that \( \Sigma_\epsilon \) and \( \Sigma \) are not correlated then the formula to compute \( \Sigma_r \) is

(2) \[ \Sigma_r = \Sigma + \Sigma_\pi \]

Formula (2) tells us that the proper relationship between the variances is \((\Sigma_r \geq \Sigma, \Sigma_\pi)\).

We can check the reference model at the boundary conditions to ensure that it is correct. In the absence of estimation error, e.g. \( \epsilon = 0 \), then \( \Sigma_r = \Sigma \). As our estimate gets worse, e.g. \( \Sigma_\pi \) increases, then \( \Sigma_r \) increases as well.

The reference model for the Black-Litterman model expected return is

(3) \[ E(r) \sim N(\pi, \Sigma_r) \]

A common misconception about the Black-Litterman reference model is that formula (1) is the reference model, and that \( \mu \) is not random. Many authors approach the problem from this point of view. Neglecting to adjust for the different reference model can cause problems.

Computing the CAPM Equilibrium Returns

As previously discussed, the prior for the Black-Litterman model is the estimated mean excess return from the CAPM equilibrium. The process of computing the CAPM equilibrium excess returns is straightforward.

CAPM is based on the concept that there is a linear relationship between risk (as measured by standard deviation of returns) and return. Further, it requires returns to be normally distributed. This model is of
the form

$$E(r)=r_f + \beta r_m + \alpha$$

Where

- $r_f$: The risk free rate.
- $r_m$: The excess return of the market portfolio.
- $\beta$: A regression coefficient computed as $\beta = \frac{\rho \sigma_p}{\sigma_m}$
- $\alpha$: The residual, or asset specific (idiosyncratic) excess return.

Under the CAPM theory the idiosyncratic risk associated with an asset's $\alpha$ is uncorrelated with the $\alpha$ from other assets and this risk can be reduced through diversification. Thus the investor is rewarded for the systematic risk measured by $\beta$, but is not rewarded for taking idiosyncratic risk associated with $\alpha$.

The Two Fund Separation Theorem, closely related to CAPM theory states that all investors should hold two assets, the CAPM market portfolio and the risk free asset. Depending on their risk aversion they will hold an arbitrary fraction of their wealth in the risky asset, and the remainder in the risk-free asset. All investors share the same risky portfolio, the CAPM market portfolio. The CAPM market portfolio is on the efficient frontier, and has the maximum Sharpe Ratio of any portfolio on the efficient frontier. Because all investors hold only this portfolio of risky assets, at equilibrium the market capitalizations of the various assets will determine their weights in the market portfolio.

Since we are starting with the market portfolio, we will be starting with a set of weights which naturally sum to 1. The market portfolio only includes risky assets, because by definition investors are rewarded only for taking on systematic risk. In the CAPM model, the risk free asset with $\beta = 0$ will not be in the market portfolio.

We will constrain the problem by asserting that the covariance matrix of the returns, $\Sigma$, is known. In practice, this covariance matrix is computed from historical return data. It could also be estimated, however there are significant issues involved in estimating a consistent covariance matrix. There is a rich body of research which claims that mean variance results are less sensitive to errors in estimating the variance and that the population covariance is more stable over time than the returns, so relying on historical covariance data does not introduce excessive model error. By computing it from actual data we know that the resulting covariance matrix will be positive definite. It is possible when estimating a covariance matrix to create one which is not positive definite, and thus not-realizable.

For the rest of this section, we will use a common notation, similar to that used in He and Litterman (1999) for all the terms in the formulas. Note that this notation is different, and conflicts, with the notation used in the section on Bayesian theory.

Here we derive the equations for 'reverse optimization' starting from the quadratic utility function

$$U = w^T \Pi - \left( \frac{\delta}{2} \right) w^T \Sigma w$$

$U$ is the investor's utility, this is the objective function during portfolio optimization.

$w$ is the vector of weights invested in each asset.
Π is the vector of equilibrium excess returns for each asset
δ is the risk aversion parameter of the market
Σ is the covariance matrix for the assets

U is a concave function, so it will have a single global maxima. If we maximize the utility with no constraints there is a closed form solution. We find the exact solution by taking the first derivative of (5) with respect to the weights (w) and setting it to 0.

\[ \frac{dU}{dw} = \Pi - \delta \Sigma w = 0 \]

Solving this for Π (the vector of excess returns) yields:

(6) \[ \Pi = \delta \Sigma w \]

In order to use formula (6) we need to have a value for δ, the risk aversion coefficient of the market. Most of the authors specify the value of δ that they used. Bevan and Winkelmann (1998) describe their process of calibrating the returns to an average Sharpe ratio based on their experience. For global fixed income (their area of expertise) they use a Sharpe ratio of 1.0. Black and Litterman (1992) use a Sharpe ratio closer to 0.5 in their example.

We can find δ by multiplying both sides of (6) by \( w^T \) and replacing vector terms with scalar terms.

(7) \[ \delta = \frac{E(r) - r_f}{\sigma^2} \]

E(r) is the total return on the market portfolio (E(r) = \( w^T \Pi + r_f \))
rf is the risk free rate
\( \sigma^2 \) is the variance of the market portfolio (\( \sigma^2 = w^T \Sigma w \))

As part of our analysis we must arrive at the terms on the right hand side of formula (7); E(r), rf, and \( \sigma^2 \) in order to calculate a value for δ. Once we have a value for δ, then we plug w, δ and Σ into formula (6) and generate the set of equilibrium asset returns. Formula (6) is the closed form solution to the reverse optimization problem for computing asset returns given an optimal mean-variance portfolio in the absence of constraints. We can rearrange formula (6) to yield the formula for the closed form calculation of the optimal portfolio weights in the absence of constraints.

(8) \[ w = (\delta \Sigma)^{-1} \Pi \]

If we feed Π, δ, and Σ back into the formula (8), we can solve for the weights (w). If we instead used historical excess returns rather than equilibrium excess returns, the results will be very sensitive to changes in Π. With the Black-Litterman model, the weight vector is less sensitive to the reverse optimized Π vector. This stability of the optimization process, is one of the strengths of the Black-Litterman model.

Herold (2005) provides insights into how implied returns can be computed in the presence of simple equality constraints such as the budget or full investment (Σw = 1) constraint.

The only missing piece is the variance of our estimate of the mean. Looking back at the reference model, we need Σπ. Black and Litterman made the simplifying assumption that the structure of the covariance matrix of the estimate is proportional to the covariance of the returns Σ. They created a parameter, τ, as the constant of proportionality. Given that assumption, Σπ = τΣ, then the prior distribution is:
This is the prior distribution for the Black-Litterman model. It represents our estimate of the mean of the distribution of excess returns.

**Specifying the Views**

This section will describe the process of specifying the investors views on the estimated mean excess returns. We define the combination of the investors views as the conditional distribution. First, by construction we will require each view to be unique and uncorrelated with the other views. This will give the conditional distribution the property that the covariance matrix will be diagonal, with all off-diagonal entries equal to 0. We constrain the problem this way in order to improve the stability of the results and to simplify the problem. Estimating the covariances between views would be even more complicated and error prone than estimating the view variances. Second, we will require views to be fully invested, either the sum of weights in a view is zero (relative view) or is one (an absolute view). We do not require a view on all assets. In addition it is actually possible for the views to conflict, the mixing process will merge the views.

We will represent the investors' $k$ views on $n$ assets using the following matrices

- $P$, a $k \times n$ matrix of the asset weights within each view. For a relative view the sum of the weights will be 0, for an absolute view the sum of the weights will be 1. Different authors compute the various weights within the view differently, Idzorek (2004) likes to use a market capitalization weighed scheme, whereas others use an equal weighted scheme.

- $Q$, a $k \times 1$ matrix of the returns for each view.

- $\Omega$, a $k \times k$ matrix of the covariance of the views. $\Omega$ is diagonal as the views are required to be independent and uncorrelated. $\Omega^{-1}$ is known as the confidence in the investor's views. The i-th diagonal element of $\Omega$ is represented as $\omega_i$.

We do not require $P$ to be invertible. Meucci (2006) describes a method of augmenting the matrices to make the $P$ matrix invertible while not changing the net results.

$\Omega$ is symmetric and zero on all non-diagonal elements, but may also be zero on the diagonal if the investor is certain of a view. This means that $\Omega$ may or may not be invertible. At a practical level we require that $\omega > 0$ so that $\Omega$ is invertible.

As an example of how these matrices would be populated we will examine some investors views. Our example will have 4 assets and two views. First, a relative view in which the investor believes that Asset 1 will outperform asset 3 by 2% with confidence $\omega_1$. Second, an absolute view in which the investor believes that asset 2 will return 3% with confidence $\omega_2$. Third, note that the investor has no view on asset 4, and thus it's return should not be directly adjusted. These views are specified as follows:

$$P = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} ; \quad Q = \begin{bmatrix} 2 \\ 3 \end{bmatrix} ; \quad \Omega = \begin{bmatrix} \omega_{11} & 0 \\ 0 & \omega_{22} \end{bmatrix}$$

Given this specification of the views we can formulate the conditional distribution mean and variance in view space as

$$P(B|A) \sim N(Q, \Omega)$$
and in asset space as

\[(10) \quad P(B|A) \sim \mathcal{N}(P^{-1}Q, [P^T\Omega^{-1}P]^{-1})\]

Remember that P may not be invertible, and even if P is invertible \([P^T\Omega^{-1}P]\) is probably not invertible, making this expression impossible to evaluate in practice. Luckily, to work with the Black-Litterman model we don't need to evaluate formula (10).

\(\Omega\), the variance of the views is inversely related to the investors confidence in the views, however the basic Black-Litterman model does not provide an intuitive way to quantify this relationship. It is up to the investor to compute the variance of the views \(\Omega\).

There are several ways to calculate \(\Omega\):

- Proportional to the variance of the prior
- Use a confidence interval
- Use the variance of residuals in a factor model
- Use Idzorek's method to specify the confidence in weight space

**Proportional to the Variance of the Prior**

We can just assume that the variance of the views will be proportional to the variance of the asset returns, just as the variance of the prior distribution is. Both He and Litterman (1999) and Meucci(2006) use this method, though they use it differently. He and Litterman (1999) set the variance of the views as follow:

\[(11) \quad \omega_i = p(\tau\Sigma)p^T \]

or

\[\Omega = \text{diag}(P(\tau\Sigma)P^T)\]

This specification of the variance, or uncertainty, of the views essentially equally weights the investor's views and the market equilibrium weights (except for the off-diagonal elements). By including \(\tau\) in the expression, the final solution becomes less dependent on the specific value of \(\tau\) selected as well. This seems to be the most common method used in the literature.

Meucci (2006) doesn't bother with the diagonalization at all, and just sets

\[\Omega = \frac{1}{c} P \Sigma P'\]

He sets \(c < 1\), and one obvious choice for \(c\) is \(\tau^{-1}\). We will see later that this form of the variance of the views lends itself to some simplifications of the Black-Litterman formulas.

**Use a Confidence Interval**

The investor can actually compute the variance of the view. This is most easily done by defining a confidence interval around the estimated mean return, e.g. Asset 2 has an estimated 3% mean return with the expectation it is 67% likely to be within the interval (2.5%,3.5%). Knowing that 67% of the normal distribution falls within 1 standard deviation of the mean, allows us to translate this into a variance of \((0.005)^2\).
Use the Variance of Residuals from a Factor Model

If the investor is using a factor model to compute the views, they can also use the variance of the residuals from the model to drive the variance of the return estimates. The general expression for a factor model of returns is:

\[ E(r) = \sum_{i=1}^{n} \beta_i f_i + \epsilon \]

Where

- \( E(r) \) is the return of the asset
- \( \beta_i \) is the factor loading for factor (i)
- \( f_i \) is the return due to factor (i)
- \( \epsilon \) is an independent normally distributed residual

And the general expression for the variance of the return from a factor model is:

\[ V(r) = BV_r(F)B^T + V(\epsilon) \]

Where

- \( B \) is the factor loading matrix
- \( F \) is the vector of returns due to the various factors

Given formula (12), and the assumption that \( \epsilon \) is independent and normally distributed, then we can compute the variance of \( \epsilon \) directly as part of the regression. While the regression might yield a full covariance matrix, the mixing model will be more robust if only the diagonal elements are used.

Beach and Orlov (2006) describe their work using GARCH style factor models to generate their views for use with the Black-Litterman model. They generate the precision of the views using the GARCH models.

Use Idzorek's Method

Idzorek (2004) describes a method for specifying the confidence in the view in terms of a percentage move of the weights on the interval from 0% confidence to 100% confidence. We will look at Idzorek's algorithm in the section on extensions.

The Estimation Model

The original Black-Litterman paper references Theil's Mixed Estimation model rather than a Bayesian estimation model, though we can get similar results from both methodologies. I chose to start with Theil's model because it is simpler and cleaner. I also work through the Bayesian version of the derivations for completeness.

With either approach, we will be estimating the mean return per the reference model, not the return distribution itself. This is important in understanding the values used for \( \tau \) and \( \Omega \), and for the computations of the variance of the prior and posterior distributions of returns.

Another way to think of the estimation model and the reference model is that while the estimated return is more accurate the variance of the distribution does not change because we have more data. The prototypical example of this would be to blend the distributions, \( P(A) \sim N(10\%, 20\%) \) and \( P(B|A) \sim N(12\%, 20\%) \). If we apply Bayes formula in a straightforward fashion, \( P(A|B) \sim N(11\%, 10\%) \). Clearly with financial data we did not really cut the variance of the return distribution in \( \frac{1}{2} \) just because
we have a slightly better estimate of the mean.

However, if the mean is the random variable, and not the distribution, then our result of \( P(A|B) \sim N(11\%, 10\%) \) makes sense. By blending these two estimates of the mean, we have an estimate of the mean with much less uncertainty (less variance) than either of the estimates.

**Theil's Mixed Estimation Model**

Theil's mixed estimation model was created for the purpose of estimating parameters from a mixture of complete prior data and partial conditional data. This is a good fit as it allows us to express views on only a subset of the asset returns, there is no requirement to express views on all of them. The views can also be expressed on a single asset, or on arbitrary combinations of the assets. The views do not even need be consistent, the estimation model will take each into account based on the investors confidence.

Theil's Mixed Estimation model starts from a linear model for the parameters to be estimated. We can use formula (1) from our reference model as a starting point.

Our simple linear model is shown below:

\[
\pi = x \beta + u
\]

Where

- \( \pi \) is the n x 1 vector of CAPM equilibrium returns for the assets.
- \( x \) is the n x n matrix \( I_n \) which are the factor loadings for our model.
- \( \beta \) is the n x 1 vector of unknown means for the asset return process.
- \( u \) is a n x n matrix of residuals from the regression where \( E(u) = 0; V(u) = E(u'u) = \Phi \) and \( \Phi \) is non-singular.

The Black-Litterman model uses a very simple linear model, the expected return for each asset is modeled by a single factor which has a coefficient of 1. Thus, \( x \), is the identity matrix. Given that \( \beta \) and \( u \) are independent, and \( x \) is constant, then we can model the variance of \( \pi \) as:

\[
V(\pi) = x' V(\beta) x + V(u)
\]

Which can be simplified to:

\[(14) \quad V(\pi) = \Sigma + \Phi\]

This ties back to formula (2) in the reference model. The total variance of the estimated return is the sum of the variance of the actual return process plus the variance of the estimate of the mean. We will come back to this relation again later in the paper.

We will pragmatically compute \( \Sigma \) from historical data for the asset returns.

Next we consider some additional information which we would like to combine with the prior. This information can be subjective views or can be derived from statistical data. We will also allow it to be incomplete, meaning that we might not have an estimate for each asset return.

\[
q = p \beta + v
\]

Where

- \( q \) is the k x 1 vector of returns for the views.
- \( p \) is the k x n vector mapping the views onto the assets.
- \( \beta \) is the n x 1 vector of unknown means for the asset return process.
is a k x k matrix of residuals from the regression where
\[ E(v) = 0; V(v) = E(v'v) = \Omega \] and \( \Omega \) is non-singular.

We can combine the prior and conditional information by writing:
\[
\begin{bmatrix}
\pi \\
q
\end{bmatrix} = \begin{bmatrix}
x \\
p
\end{bmatrix} \hat{\beta} + \begin{bmatrix}
u \\
v
\end{bmatrix}
\]

Where the expected value of the residual is 0, and the expected value of the variance of the residual is
\[
V\left( \begin{bmatrix} u \\ v \end{bmatrix} \right) = E\left( \begin{bmatrix} u \\ v \end{bmatrix} u' v' \right) = \begin{bmatrix} \Phi & 0 \\ 0 & \Omega \end{bmatrix}
\]

We can then apply the generalized least squares procedure, which leads to estimating \( \hat{\beta} \) as
\[
\hat{\beta} = \begin{bmatrix} x' & p' \end{bmatrix} \begin{bmatrix} \Phi & 0 \\ 0 & \Omega \end{bmatrix}^{-1} \begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} x' & p' \end{bmatrix} \begin{bmatrix} \Phi & 0 \\ 0 & \Omega \end{bmatrix}^{-1} \begin{bmatrix} \pi \\ q \end{bmatrix}
\]

This can be rewritten (dropping x as it is the identity matrix) without the matrix notation as
\[
\hat{\beta} = \left( \Phi^{-1} + p' \Omega^{-1} p \right)^{-1} \left( \Phi^{-1} \pi + p' \Omega^{-1} q \right)
\]

This new \( \hat{\beta} \) is the weighted average of the estimates, where the weighting factor is the inverse of the variance of the estimate. It is also the best linear unbiased estimate given the data, and has the property that it minimizes the variance of the residual. Note that given a new \( \hat{\beta} \), we also should have an updated expectation for the variance of the residual.

We can reformulate our combined relationship in terms of our estimate of \( \hat{\beta} \) and a new residual \( \tilde{u} \) as
\[
\begin{bmatrix}
\pi \\
q
\end{bmatrix} = \begin{bmatrix}
x \\
p
\end{bmatrix} \hat{\beta} + \tilde{u}
\]

Once again \( E(\tilde{u}) = 0 \), so we can derive the expression for the variance of the new residual as:
\[
V(\tilde{u}) = E(\tilde{u}'\tilde{u}) = \left[ \Phi^{-1} + r' \Psi^{-1} r \right]^{-1}
\]

and the total variance is
\[
V(\begin{bmatrix} y & \pi \end{bmatrix}) = V(\hat{\beta}) + V(\tilde{u})
\]

We began this section by asserting that the variance of the return process is a known quantity derived from historical estimates. Improved estimation of the quantity \( \hat{\beta} \) does not change our estimate of the variance of the return distribution, \( \Sigma \). Because of our improved estimate, we do expect that the variance of the estimate (residual) has decreased, thus the total variance has changed. We can simplify the variance formula (14) to
\[
V(\begin{bmatrix} y & \pi \end{bmatrix}) = \Sigma + V(\tilde{u})
\]

This is a clearly intuitive result, consistent with the realities of financial time series. We have combined two estimates of the mean of a distribution to arrive at a better estimate of the mean. The variance of this estimate has been reduced, but the actual variance of the underlying process has not
changed. Given our uncertain estimate of the process, the total variance of our estimated process has also improved incrementally, but it has the asymptotic limit that it cannot be less than the variance of the actual underlying process.

Appendix A contains a more detailed derivation of formulas (15) and (16).

**A Quick Introduction to Bayes Theory**

This section provides a quick overview of the relevant portion of Bayes theory in order to create a common vocabulary which can be used in analyzing the Black-Litterman model from a Bayesian point of view.

Bayes theory states

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\]

- **P(A|B)**: The conditional (or joint) probability of A, given B. Also known as the posterior distribution. We will call this the posterior distribution from here on.
- **P(B|A)**: The conditional probability of B given A. Also known as the sampling distribution. We will call this the conditional distribution from here on.
- **P(A)**: The probability of A. Also known as the prior distribution. We will call this the prior distribution from here on.
- **P(B)**: The probability of B. Also known as the normalizing constant.

When actually applying this formula and solving for the posterior distribution, the normalizing constant will disappear into the constants of integration so from this point on we will ignore it.

A general problem in using Bayes theory is to identify an intuitive and tractable prior distribution. One of the core assumptions of the Black-Litterman model (and Mean-Variance optimization) is that asset returns are normally distributed. For that reason we will confine ourselves to the case of normally distributed conditional and prior distributions. Given that the inputs are normal distributions, then it follows that the posterior will also be normally distributed. When the prior distribution and the posterior have the same structure, the prior is known as a conjugate prior. Given interest there is nothing to keep us from building variants of the Black-Litterman model using different distributions, however the normal distribution is generally the most straightforward.

Another core assumption of the Black-Litterman model is that the variance of the prior and the conditional distributions about the actual mean are known, but the actual mean is not known. This case, known as “Unknown Mean and Known Variance” is well documented in the Bayesian literature. This matches the model which Theil uses where we have an uncertain estimate of the mean, but know the variance.

We define the significant distributions below:

The prior distribution

\[
P(A) \sim N(x, S/n)
\]

where S is the sample variance of the distribution about the mean, with n samples then S/n is the variance of the estimate of x about the mean.
The conditional distribution

(20) \( P(B|A) \sim N(\mu,\Omega) \)

\( \Omega \) is the uncertainty in the estimate \( \mu \) of the mean, it is not the variance of the distribution about the mean.

Then the posterior distribution is specified by

(21) \( P(A|B) \sim N([\Omega^{-1}\mu + nS^{-1}x]^{T}[\Omega^{-1} + nS^{-1}]^{-1},(\Omega^{-1} + nS^{-1})^{-1}) \)

The variance term in (21) is the variance of the estimated mean about the actual mean.

In Bayesian statistics the inverse of the variance is known as the precision. We can describe the posterior mean as the weighted mean of the prior and conditional means, where the weighting factor is the respective precision. Further, the posterior precision is the sum of the prior and conditional precision. Formula (21) requires that the precisions of the prior and conditional both be non-infinite, and that the sum is non-zero. Infinite precision corresponds to a variance of 0, or absolute confidence. Zero precision corresponds to infinite variance, or total uncertainty.

A full derivation of formula (21). using the PDF based Bayesian approach is shown in Appendix B.

As a first check on the formulas we can test the boundary conditions to see if they agree with our intuition. If we examine formula (21) in the absence of a conditional distribution, it should collapse into the prior distribution.

\( \sim N([nS^{-1}x][nS^{-1}]^{-1},(nS^{-1})^{-1}) \)

(22) \( \sim N(x,S/n) \)

As we can see in formula (22), it does indeed collapse to the prior distribution. Another important scenario is the case of 100% certainty of the conditional distribution, where \( S \), or some portion of it is 0, and thus \( S \) is not invertible. We can transform the returns and variance from formula (21) into a form which is more easy to work with in the 100% certainty case.

(23) \( P(A|B) \sim N(x + (S/n)(\frac{A^{-1} + B^{-1})^{-1} - A - A(A+B)^{-1}A}{(S/n)(\Omega + S/n)^{-1}(S/n)}) \)

This transformation relies on the result that \( (A^{-1} + B^{-1})^{-1} = A - A(A+B)^{-1}A \). It is easy to see that when \( S \) is 0 (100% confidence in the views) then the posterior variance will be 0. If \( \Omega \) is positive infinity (the confidence in the views is 0%) then the posterior variance will be \( (S/n) \).

We will revisit equations (21) and (23) later in this paper where we transform these basic equations into the various parts of the Black-Litterman model. Appendices C and D contain derivations of the alternate Black-Litterman formulas from the standard form, analogous to the transformation from (21) to (23).

The final piece of the puzzle is the total variance of the posterior returns. From our reference model, formula (2), we know we need to add the posterior variance of the estimated mean to the 'known covariance' of the actual returns to arrive at the total covariance we will use as input to the mean-variance optimizer.

He and Litterman (1999) adopt this convention and compute the variance of the posterior returns by adding the sample variance to the variance of the posterior distribution as in

(24) \( S_p = S + M \)
where $M$ is the variance of the posterior distribution about the actual mean and $S$ is the 'known covariance' of returns.

In the absence of views, formula (24) simplifies to

$$M_p = \frac{(n+1)}{n}S$$

**Using Bayes Theorem for the Estimation Model**

One of the major assumptions made by the Black-Litterman model is that the covariance of the estimated mean is proportional to the covariance of the actual returns. The parameter $\tau$ will serve as the constant of proportionality. It takes the place of $1/n$ in formula (19). The new prior is:

(25) \hspace{1cm} P(A) \sim N(\Pi, \tau S)

This is the prior distribution for the Black-Litterman model.

We can now apply Bayes theory to the problem of blending the prior and conditional distributions to create a new posterior distribution of the asset returns. Given equations (21), (25) and (10) we can apply Bayes Theorem and derive our formula for the posterior distribution of asset returns.

Substituting (25) and (10) into (21) we have the following distribution

(26) \hspace{1cm} P(A|B) \sim N[(\tau \Sigma)^{-1} \Pi + P^T \Omega^{-1} Q] \frac{1}{((\tau \Sigma)^{-1} + P^T \Omega^{-1} P)^{-1}}.

This is sometimes referred to as the Black-Litterman master formula. A complete derivation of the formula is shown in Appendix B. An alternate representation of the same formula for the mean returns ($E(r)$) and covariance ($M$) is

(27) \hspace{1cm} E(r) = \Pi + \tau \Sigma P^T [(P \tau \Sigma P^T) + \Omega]^{-1} [Q - \Pi \Pi^T]

(28) \hspace{1cm} M = ((\tau \Sigma)^{-1} + P^T \Omega^{-1} P)^{-1}

The derivation of formula (27) is shown in Appendix D. Remember that $M$, the posterior variance, is the variance of the posterior mean estimate about the actual mean. It is the uncertainty in the posterior mean estimate, and is not the variance of the returns. In order to compute the variance of the returns so that we can use it in a mean-variance optimizer we need to apply formula (24). This is mentioned in He and Litterman (1999) but not in any of the other papers.

(29) \hspace{1cm} \Sigma_p = \Sigma + M

Substituting the posterior variance from (28) we get

$$\Sigma_p = \Sigma + ((\tau \Sigma)^{-1} + P^T \Omega^{-1} P)^{-1}$$

In the absence of views this reduces to

(30) \hspace{1cm} \Sigma_p = \Sigma + (\tau \Sigma) = (1 + \tau)\Sigma

Thus when applying the Black-Litterman model in the absence of views the variance of the estimated returns will be greater than the prior distribution variance. We see the impact of this formula in the results shown in He and Litterman (1999). In their results, the investor's weights sum to less than 1 if they have no views.\(^3\) Idzorek (2004) and most other authors do not compute a new posterior variance, but instead use the known input variance of the returns about the mean. Several of the authors set $\tau = 1$ along with using the variance of returns. We will consider these different reference models in a later

---

\(^3\) This is shown in table 4 and mentioned on page 11 of He and Litterman (1999).
section.

Once we dig into the topic a little more we realize that if we have only partial views, views on some assets, then by using a posterior estimate of the variance we will tilt the posterior weights towards assets with lower variance (higher precision of the estimated mean) and away from assets with higher variance (lower precision of the estimated mean). Thus the existence of the views and the updated covariance will tilt the optimizer towards using or not using those assets. This tilt will not be very large if we are working with a small value of $\tau$, but it will be measurable.

Since we are building the covariance matrix, $\Sigma$, from historical data we can compute $\tau$ from the number of samples. We can also estimate $\tau$ based on our confidence in the prior distribution. Note that both of these techniques provide some intuition for selecting a value of $\tau$ which is closer to 0 than to 1. Black and Litterman (1992), He and Litterman (1999) and Idzorek (2004) all indicate that in their calculations they used small values of $\tau$, on the order of 0.025 – 0.050. Satchell and Scowcroft (2000) state that many investors use a $\tau$ around 1 which does not seem to have any intuitive connection here.

We can check our results by seeing if the results match our intuition at the boundary conditions. Given formula (27) it is easy to let $\Omega \to 0$ showing that the return under 100% certainty of the views is

$$E(r) = \Pi + \Sigma^T [P\Sigma^T]^{-1}[Q - P\Pi]$$

Thus under 100% certainty of the views, the estimated return is insensitive to the value of $\tau$ used.

Furthermore, if $P$ is invertible which means that it we have also offered a view on every asset, then

$$E(r) = P^{-1}Q$$

If the investor is not sure about their views, so $\Omega \to \infty$, then formula (27) reduces to

$$E(r) = \Pi$$

Finding an analytically tractable way to express and compute the posterior variance under 100% certainty is a challenging problem. Formula (21) above works only if $(P^T\Omega^{-1}P)$ is invertible which is not usually the case because the posterior variance in asset space is also not usually tractable.

The alternate formula for the posterior variance derived from (23) is

$$M = \tau \Sigma - \tau \Sigma^T [P\tau \Sigma^T + \Omega]^{-1} \Sigma \tau$$

If $\Omega \to 0$ (total confidence in views, and every asset is in at least one view) then formula (32) can be reduced to $M = 0$. If on the other hand the investor is not confident in their views, $\Omega \to \infty$, then formula (32) can be reduced to $M = \tau \Sigma$.

[Meucci, 2005] describes the transformation from (23) to (32), but does not show the full derivation. I have included that derivation in Appendix B.

**Calibrating $\tau$**

This section will discuss some empirical ways to select and calibrate the value of $\tau$.

The first method to calibrate $\tau$ relies on falling back to basic statistics. When estimating the mean of a distribution the uncertainty (variance) of the mean estimate will be proportional to the number of samples. Given that we are estimating the covariance matrix from historical data, then
Given that we usually aim for a number of samples around 60 (5 years of monthly samples) then $\tau$ is on the order of 0.02. This is consistent with several of the papers which indicate they used values of $\tau$ on the range (0.025, 0.05). This is probably also consistent with the model to compute the variance of the distribution, $\Sigma$.

We can also calibrate $\tau$ to the amount invested in the risk free asset given the prior distribution. Here we see that the portfolio invested in risky assets given the prior views will be

$$w = \Pi \left[ \delta \left( 1 + \tau \right) \Sigma \right]^{-1}$$

Thus the weights allocated to the assets are smaller by $1/(1+\tau)$ than the CAPM market weights. This is because our Bayesian investor is uncertain in their estimate of the prior, and they do not want to be 100% invested in risky assets.

**The Alternative Reference Model**

This section will discuss the most common alternative reference model used with the Black-Litterman estimation model.

The most common alternative reference model is the one used in Satchell and Scowcroft (2000).

$$E(r) \sim N(\mu, \Sigma)$$

In this reference model, $\mu$ is normally distributed with variance $\Sigma$. We estimate $\mu$, but it is not considered a random variable. This is commonly described as having a $\tau = 1$, but because of the reference model it eliminates $\tau$ as a parameter. In this model $\Omega$ becomes the covariance of the views around the investor's estimate of the mean return, just as $\Sigma$ is the covariance of the prior about it's mean. Given these specifications for the covariances of the prior and conditional, it is infeasible to have an updated covariance for the posterior. For example, it would require that the covariance of the posterior is $1/2$ the covariance of the prior and conditional if they had equal covariances. This conflicts with our earlier statements that higher precision in our estimate of the mean return doesn't cause the covariance of the return distribution to shrink by the same amount.

The primary artifacts of this new reference model are, first $\tau = 1$, or is non-existant, and second, the investor's portfolio weights in the absence of views equal the CAPM portfolio weights. Finally at implementation time there is no need or use of formulas (28) or (29).

Note that none of the authors prior to Meucci (2008) except for Black and Litterman (1992), and He and Litterman (1999) make any mention of the details of the Black-Litterman reference model, or of the different reference model used by most authors. It is unclear to this author why this has occurred.

In the Black-Litterman reference model, the updated posterior variance of the mean estimate will be lower than either the prior or conditional variance of the mean estimate, indicating that the addition of more information will reduce the uncertainty of the model. The variance of the returns from formula (29) will never be less than the prior variance of returns. This matches our intuition as adding more information should reduce the uncertainty of the estimates. Given that there is some uncertainty in this value (M), then formula (29) provides a better estimator of the variance of returns than the prior variance of returns.
**The Impact of τ**

The meaning and impact of the parameter τ causes a great deal of confusion for many users of the Black-Litterman model. In the literature we appear see authors divided into two groups over τ. In reality, τ is not the significant difference, the reference model is the difference and the author's specification of τ just an artifact of the reference model they use.

The first group thinks τ should be a small number on the order of 0.025-0.05, and includes He and Litterman (1999), Black and Litterman (1992) and Idzorek(2004). The second group thinks τ should be near 1 and includes Satchell and Scowcroft (2000), Meucci (2005) and others.

A better division of the authors has to do with the reference model, the first group except for Idzorek (2004) uses the Black-Litterman reference model. The second group uses the alternative reference model described above. Idzorek (2004) introduces a new method for specifying the investors confidence in the views and τ does not impact his formulas. Satchell and Scowcroft (2000) describe a model with a stochastic τ, but this is really stochastic variance of returns.

Given the Black-Litterman reference model we can still perform an exercise to understand the impact of τ on the results. We will start with the expression for Ω used by He and Litterman (1999).

\[ \Omega = P (\tau \Sigma) P^T \]

If we assume P is of full rank, then we can substitute this into formula (15) as

\[ E(r) = \Pi + \tau \Sigma P^T [(\tau \Sigma P^T) + \Omega]^{-1} [Q - \Pi \Pi] \]

\[ = \Pi + \tau \Sigma P^T [(\tau \Sigma P^T) + \tau \Sigma (\tau \Sigma P^T)^{-1} P [Q - \Pi \Pi] \]

\[ = \Pi + \tau \Sigma P^T [2 (\tau \Sigma P^T)^{-1} (\tau \Sigma)^{-1} P [Q - \Pi \Pi] \]

\[ = \Pi + (1/2) \tau \Sigma P^T [Q - \Pi \Pi] \]

\[ = \Pi + (1/2) \tau \Sigma P^T [Q - \Pi] \]

Clearly using formula (33) is just a simplification and does not do justice to investors views, but we can still see that setting Ω proportional to τ, will eliminate τ from the final formula for E(r).

We can see a similar result if we substitute formula (33) into formula (32).

\[ M = \tau \Sigma - \tau \Sigma P^T [\tau \Sigma P^T + \Omega]^{-1} \tau \Sigma \]

\[ = \tau \Sigma - \tau \Sigma P^T [(\tau \Sigma P^T) + \tau \Sigma (\tau \Sigma P^T)^{-1} P \tau \Sigma \]

\[ = \tau \Sigma - 2 \tau \Sigma P^T [2 (\tau \Sigma P^T)^{-1} (\tau \Sigma)^{-1} P \tau \Sigma \]

\[ = \tau \Sigma - (1/2) \tau \Sigma P^T [\tau \Sigma)^{-1} (P^T)^{-1} P \tau \Sigma \]

\[ = \tau \Sigma - (1/2) (\tau \Sigma) \]

\[ = (1/2) (\tau \Sigma) \]

Note that τ is not eliminated from formula (35). We can also observe that if τ is on the order of 1 and were to use formula (29) that the uncertainty in the estimate of the mean would be a significant fraction of the variance of the returns. With the alternative reference model, no posterior variance calculations are performed and the mixing is weighted by the variance of returns.

In both cases, our choice for Ω has evenly weighted the prior and conditional distributions in the estimation of the posterior distribution. This matches our intuition when we consider we have blended two inputs, for both of which we have the same level of uncertainty. The posterior distribution will be the average of the two distributions.

If we instead solve for the more useful general case of \( \Omega = \alpha P (\tau \Sigma) P^T \) where \( \alpha \geq 1 \), substituting into

© Copyright 2008, Jay Walters
(27) and following the same logic as used to derive (35) we get

\[E(\pi) = \Pi + \frac{1}{1+\alpha}[P^{-1}Q - \Pi]\]

This parameterization of the uncertainty is specified in Meucci (2005) and it allows us an option between using the same uncertainty for the prior and views, and having to specify a separate and unique uncertainty for each view. Given that we are essentially multiplying the prior covariance matrix by a constant this parameterization of the uncertainty of the views does not have a negative impact on the stability of the results.

In summary, if the investor uses the alternative reference model and makes \(\Omega\) proportional to \(\tau\Sigma\), then they need only calibrate the constant of proportionality, \(\alpha\), which indicates their relative confidence in their views versus the equilibrium. If they use the Black-Litterman reference model and set \(\Omega\) proportional to \(\tau\Sigma\), then return estimate will not depend on the value of \(\tau\), but the posterior covariance of returns will depend on the proper calibration of \(\tau\).

**An Asset Allocation Process**

The Black-Litterman model is just one part of an asset allocation process. Bevan and Winkelmann (1998) document the asset allocation process they use in the Fixed Income Group at Goldman Sachs. At a minimum, a Black-Litterman oriented investment process would have the following steps:

- Determine which assets constitute the market
- Compute the historical covariance matrix for the assets
- Use reverse optimization to compute the CAPM equilibrium returns for the assets
- Specify views on the market
- Blend the CAPM equilibrium returns with the views using the Black-Litterman model
- Feed the estimates (estimated returns, covariances) generated by the Black-Litterman model into a portfolio optimizer.
- Select the efficient portfolio which matches the investors risk preferences

A further discussion of each step is provided below.

The first step is to determine the scope of the market. For an asset allocation exercise this would be identifying the individual asset classes to be considered. For each asset class the weight of the asset class in the market portfolio is required. Then a suitable proxy return series for the excess returns of the asset class is required. Between these two requirements it can be very difficult to integrate illiquid asset classes such as private equity or real estate into the model.

Once the proxy return series have been identified, and returns in excess of the risk free rate have been calculated, then a covariance matrix can be calculated. Typically the covariance matrix is calculated from the highest frequency data available, e.g. daily, and then scaled up to the appropriate time frame. Investor's often use an exponential weighting scheme to provide more emphasis to more recent data and less to older data.

Now we can run a reverse optimization on the market portfolio to compute the equilibrium excess returns for each asset class. Part of this step includes computing a \(\delta\) value for the market portfolio. This can be calculated from the return and standard deviation of the market portfolio. Bevan and Winkelmann (1998) discuss the use of an expected Sharpe Ratio target for the calibration of \(\delta\). For
their international fixed income investments they used an expect Sharpe Ratio of 1.0 for the market. The investor then needs to calibrate $\tau$ in some manner. This value is usually on the order of $0.025 \sim 0.050$.

At this point almost all of the machinery is in place. Next the investor needs to specify views on the market. These views can impact one or more assets, in any combination. They need to specify the assets involved in each view, the absolute or relative return of the view, and their uncertainty in the return for the view as measured by a variance around the actual mean (estimation error).

Appendix E shows the process of cranking through formulas (27), (32) and (29) to compute the new posterior estimate of the returns and the covariance of the posterior returns. These values will be the inputs to some type of optimizer, a mean-variance optimizer being the most common. If the user generates the optimal portfolios for a series of returns, then they can plot an efficient frontier.

**Results**

This section of the document will step through a comparison the results of the various authors. The java programs used to compute these results are all available as part of the akutan open source finance project at sourceforge.net. All of the mathematical functions were built using the Colt open source numerics library for Java. Any small differences between my results and the authors reported results are most likely the result of rounding of inputs and/or results.

When reporting results most authors have just reported the portfolio weights from an unconstrained optimization using the posterior mean and variance. Given that the vector $\Pi$ is the excess return vector, then we do not need a budget constraint ($\Sigma w_i = 1$) as we can safely assume any 'missing' weight is invested in the risk free asset which has expected return 0 and variance 0. This calculation comes from formula (8).

$$w = \Pi (\delta \Sigma_p)^{-1}$$

As a first test of our algorithm we verify that when the investor has no views that the weights are correct, substituting formula (30) into (8) we get

$$w_{nv} = \Pi (\delta (1+\tau) \Sigma)^{-1}$$

(37) $w_{nv} = w / (1+\tau)$

Given this result, it is clear that the output weights with no views will be impacted by the choice of $\tau$ when the Black-Litterman reference model is used. He and Litterman (1999) indicate that if our investor is a Bayesian, then they will not be certain of the prior distribution and thus would not be fully invested in the risky portfolio at the start. This is consistent with formula (37).

**Matching the Results of He and Litterman**

First we will consider the results shown in He and Litterman (1999). These results are the easiest to reproduce and they seem to stay close to the model described in the original Black and Litterman (1992) paper. As Robert Litterman co-authored both this paper and the original paper, it makes sense they would be consistent.

He and Litterman (1999) set $\Omega = \text{diag}(P^T \tau \Sigma P)$ essentially making the uncertainty of the views equivalent to the uncertainty of the equilibrium estimates. They select a small value for $\tau$ (0.05), and they use the Black-Litterman reference model and the updated posterior variance of returns as

© Copyright 2008, Jay Walters
calculated in formulas (32) and (29).

Table 1 – These results correspond to Table 7 in [He and Litterman, 1999].

<table>
<thead>
<tr>
<th>Asset</th>
<th>P₀</th>
<th>P₁</th>
<th>μ</th>
<th>wₑᵥ/(1+τ)</th>
<th>w*</th>
<th>w* - wₑᵥ/(1+τ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>0.0</td>
<td>0.0</td>
<td>4.3</td>
<td>16.4</td>
<td>1.5%</td>
<td>.0%</td>
</tr>
<tr>
<td>Canada</td>
<td>0.0</td>
<td>1.0</td>
<td>8.9</td>
<td>2.1%</td>
<td>53.9%</td>
<td>51.8%</td>
</tr>
<tr>
<td>France</td>
<td>-0.295</td>
<td>0.0</td>
<td>9.3</td>
<td>5.0%</td>
<td>-.5%</td>
<td>-5.4%</td>
</tr>
<tr>
<td>Germany</td>
<td>1.0</td>
<td>0.0</td>
<td>10.6</td>
<td>5.2%</td>
<td>23.6%</td>
<td>18.4%</td>
</tr>
<tr>
<td>Japan</td>
<td>0.0</td>
<td>0.0</td>
<td>4.6</td>
<td>11.0%</td>
<td>11.0%</td>
<td>.0%</td>
</tr>
<tr>
<td>UK</td>
<td>-0.705</td>
<td>0.0</td>
<td>6.9</td>
<td>11.8%</td>
<td>-1.1%</td>
<td>-13.0%</td>
</tr>
<tr>
<td>USA</td>
<td>0.0</td>
<td>-1.0</td>
<td>7.1</td>
<td>58.6%</td>
<td>6.8%</td>
<td>-51.8%</td>
</tr>
<tr>
<td>q</td>
<td>5.0</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ω/τ</td>
<td>.043</td>
<td>.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ</td>
<td>.193</td>
<td>.544</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 contains results computed using the akutan implementation of Black-Litterman and the input data for the equilibrium case and the investor's views from He and Litterman (1999). The values shown for w* exactly match the values shown in their paper.

**Matching the Results of Idzorek**

This section of the document describes the efforts to reproduce the results of Idzorek (2004). In trying to match Idzorek's results I found that he used the alternative reference model, which leaves Σ, the known variance of the returns from the prior distribution, as the variance of the posterior returns. This is a significant difference from the algorithm used in He and Litterman (1999), but in the end given that Idzorek used a small value of τ, the differences amounted to approximately 50 basis points per asset. Tables 2 and 3 below illustrate computed results with the data from his paper and how the results differ between the two versions of the model.

Table 2 contains results generated using the data from Idzorek (2004) and the Black-Litterman model as described by He and Litterman (1999). Table 3 shows the same results as generated by the alternative reference model variant of the algorithm. This alternative model variant appears to be the method used by Idzorek.
Table 2 – He and Litterman version of Black-Litterman Model with Idzorek data (Corresponds with data in Idzorek's Table 6).

<table>
<thead>
<tr>
<th>Asset Class</th>
<th>$\mu$</th>
<th>$w_{eq}$</th>
<th>$w^*$</th>
<th>Black-Litterman Reference Model</th>
<th>Idzorek's Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Bonds</td>
<td>.07</td>
<td>18.87%</td>
<td>28.96%</td>
<td>10.09%</td>
<td>10.54</td>
</tr>
<tr>
<td>Intl Bonds</td>
<td>.50</td>
<td>25.49%</td>
<td>15.41%</td>
<td>-10.09%</td>
<td>-10.54</td>
</tr>
<tr>
<td>US LG</td>
<td>6.50</td>
<td>11.80%</td>
<td>9.27%</td>
<td>-2.52%</td>
<td>-2.73</td>
</tr>
<tr>
<td>US LV</td>
<td>4.33</td>
<td>11.80%</td>
<td>14.32%</td>
<td>2.52%</td>
<td>-2.73</td>
</tr>
<tr>
<td>US SG</td>
<td>7.55</td>
<td>1.31%</td>
<td>1.03%</td>
<td>-.28%</td>
<td>-0.30</td>
</tr>
<tr>
<td>US SV</td>
<td>3.94</td>
<td>1.31%</td>
<td>1.59%</td>
<td>.28%</td>
<td>0.30</td>
</tr>
<tr>
<td>Intl Dev</td>
<td>4.94</td>
<td>23.59%</td>
<td>27.74%</td>
<td>4.15%</td>
<td>3.63</td>
</tr>
<tr>
<td>Intl Emg</td>
<td>6.84</td>
<td>3.40%</td>
<td>3.40%</td>
<td>.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

Note that the results in Table 2 are close, but for several of the assets the difference is about 50 basis points. The values shown in Table 3 are within 4 basis points, essentially matching the results reported by Idzorek.

Table 3 – Alternative Reference Model version of the Black-Litterman Model with Idzorek data (Corresponds with data in Idzorek's Table 6).

<table>
<thead>
<tr>
<th>Country</th>
<th>$\mu$</th>
<th>$w_{eq}$</th>
<th>$w$</th>
<th>Alternative Reference Model</th>
<th>Idzorek's Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Bonds</td>
<td>.07</td>
<td>19.34%</td>
<td>29.89%</td>
<td>10.55%</td>
<td>10.54</td>
</tr>
<tr>
<td>Intl Bonds</td>
<td>.50</td>
<td>26.13%</td>
<td>15.58%</td>
<td>-10.55%</td>
<td>-10.54</td>
</tr>
<tr>
<td>US LG</td>
<td>6.50</td>
<td>12.09%</td>
<td>9.37%</td>
<td>-2.72%</td>
<td>-2.73</td>
</tr>
<tr>
<td>US LV</td>
<td>4.33</td>
<td>12.09%</td>
<td>14.81%</td>
<td>2.72%</td>
<td>-2.73</td>
</tr>
<tr>
<td>US SG</td>
<td>7.55</td>
<td>1.34%</td>
<td>1.04%</td>
<td>-.30%</td>
<td>-0.30</td>
</tr>
<tr>
<td>US SV</td>
<td>3.94</td>
<td>1.34%</td>
<td>1.64%</td>
<td>.30%</td>
<td>0.30</td>
</tr>
<tr>
<td>Intl Dev</td>
<td>4.94</td>
<td>24.18%</td>
<td>27.77%</td>
<td>3.59%</td>
<td>3.63</td>
</tr>
<tr>
<td>Intl Emg</td>
<td>6.84</td>
<td>3.49%</td>
<td>3.49%</td>
<td>.0%</td>
<td>0</td>
</tr>
</tbody>
</table>

© Copyright 2008, Jay Walters
On Mankert's Sampling Theoretic Analysis

Mankert (2006) derives the Black-Litterman 'master formula' using a Sampling Theory approach which yields $\tau$ as a ratio of the confidence in the prior to the confidence in the sampling distribution. She also provides a full derivation of formula (27) from formula (26).

Her argument about $\tau$ is interesting, and given that most authors use the alternative reference model, and that $\Omega$ is often set proportional to $P(\Sigma)P^T$ it is not unreasonable. If we attempt to perform the same manipulations to the Black-Litterman reference model we will see her thesis does not hold.

Given the derivation of the posterior distribution created by the mixing of a normally distributed statistical prior and a subjective conditional shown in Appendix A, we can by replacing the subject conditional with a statistical conditional distribution arrive at formula (21) by the same logic.

Given

$$P(A) \sim N\left( x_p \frac{S_p}{n} \right), \text{ given } n \text{ samples}$$

$$P(B|A) \sim N\left( x_c \frac{S_c}{m} \right), \text{ given } m \text{ samples}$$

then

$$P(A|B) \sim N\left( \left[ \left( S_p \frac{m}{n} \right)^{-1} + \left( S_c \frac{m}{n} \right)^{-1} \right]^{-1} \left[ x_p \left( \frac{S_p}{n} \right)^{-1} + x_c \left( \frac{S_c}{m} \right)^{-1} \right], \left[ \left( S_p \frac{m}{n} \right)^{-1} + \left( S_c \frac{m}{n} \right)^{-1} \right]^{-1} \right)$$

Next we want to attempt to move the (m) terms out of each term so we can eliminate them yielding:

$$P(A|B) \sim N\left( m \left[ \left( S_p \frac{m}{n} \right)^{-1} + (S_c)^{-1} \right]^{-1} \left[ m \left[ x_p \left( S_p \frac{m}{n} \right)^{-1} + x_c \left( S_c \right)^{-1} \right]\right], m \left[ \left( S_p \frac{m}{n} \right)^{-1} + (S_c)^{-1} \right]^{-1} \right)$$

In the mean term the m's cancel out and we see the formula Mankert used to support her thesis

$$\left[ \left( S_p \frac{m}{n} \right)^{-1} + (S_c)^{-1} \right]^{-1} \left[ x_p \left( S_p \frac{m}{n} \right)^{-1} + x_c \left( S_c \right)^{-1} \right]$$

If $\Omega = S_c$ then we can set $\tau = m/n$, the ratio of uncertainty in the views to the uncertainty in the prior.

However in the posterior variance of the estimate an extra m term remains, this makes the covariance term

$$\left( \frac{1}{m} \right) \left[ \left( \frac{m}{n} S_p \right)^{-1} \right]$$

Thus if we neglect the variance of the posterior estimate (as is done in the alternative reference model) the algebra works.

There is one remaining issue, and that is $\Omega = S_c$ must be the variance of the return distributions for the views. In the Black-Litterman reference model, $\Omega$ is the uncertainty of the views in regards to our estimate of the mean. That is, $\Omega = S_c/m$, and we do not separately estimate $S_c$ or $m$. In the alternative reference model, $\tau$ is either 1 or not in the model, which means that essentially $\Sigma = S_p$ and so $\Omega = S_c$.

Again, Mankert's thesis is consistent with the alternative reference model, but not the Black-Litterman reference model.
**Additional Work**

This section provides a brief discussion of efforts to reproduce results from some of the other research papers.

Of the major papers on the Black-Litterman model, there are two which would be very useful to reproduce, Satchell and Scowcroft (2000) and Black and Litterman (1992). Satchell and Scowcroft (2000) does not provide enough data in their paper to reproduce their results. They have several examples, one with 11 countries equity returns plus currency returns, and one with fifteen countries. They don't provide the covariance matrix for either example, and so their analysis cannot be reproduced. It would be interesting to confirm that they use the alternative reference model by reproducing their results.

Black and Litterman (1992) do provide what seems to be all the inputs to their analysis, however they chose a non-trivial example including partially hedged equity and bond returns. This requires the application of some constraints to the reverse optimization process which I have been unable to formulate as of this time. I plan on continuing this work with the goal of verifying the details of the Black-Litterman implementation used by Black and Litterman.

**Extensions to the Black-Litterman Model**

In this section I will cover the extensions to the Black-Litterman model proposed in Idzorek (2004), Fusai and Meucci (2003) and Krishnan and Mains (2006).

Idzorek (2004) presents a means to calibrate the confidence or variance of the investors views in a simple and straightforward method. Fusai and Meucci (2003) propose a way to measure how consistent a posterior estimate of the mean is with regards to the prior, or some other estimate. Krishnan and Mains (2006) present a method to incorporate additional factors into the model.

**Idzorek’s Extension**

Idzorek's extension is to specify the investors confidence as a percentage (0-100%) where the confidence measures the change in weight of the posterior from the prior estimate (0%) to the conditional estimate (100%). This linear relation is shown below

\[
\text{confidence} = \frac{(\hat{w} - w_{mkt})}{(w_{100} - w_{mkt})}
\]

The extension proposed by Idzorek is very useful, and reduces the complexity of the model for non-quantitative users. It's primary goal is to allow the specification of investor view confidence as a percentage, where the model will back out the proper value of \( \Omega \) to reach the confidence level specified. A side effect of this method is that it is insensitive to the choice of \( \tau \).

Basically, what is done, is a linear model for the values \( \omega_i \) is specified. We can consider the value of the unconstrained weights under no views to be equivalent to having a view with 0% confidence, and for the other boundary point we can use the unconstrained weights given 100% certainty in the view. The formula below illustrates the model

\[
(39) \quad \text{confidence} = \frac{(\hat{w} - w_{mkt})}{(w_{100} - w_{mkt})}
\]

- \( w_{100} \) is the weight of the asset under 100% certainty
- \( w_{mkt} \) is the weight of the asset under no views
w is the weight of the asset under the specified view.

We can solve for the view confidence analytically if we take a few shortcuts. First we will use the following form of the uncertainty of the views

\[ \Omega = \alpha P \Sigma P^T \]

Note that \( \Omega \) is a 1x1 matrix for a single view, and thus it is equivalent whether it is full or only on the diagonal. This allows us to substitute formula (36) into formula (8).

\[ \hat{w} = E(r)(\delta \Sigma)^{-1} \]

\[ \hat{w} = \left[ \Pi + \frac{1}{(1+\alpha)} \left[ P^{-1} Q - \Pi \right] \right] (\delta \Sigma)^{-1} \]

Now we can solve formula (8) at the boundary conditions, \( w_{mk} \) and \( w_{100} \).

\[ w_{mk} = \Pi (\delta \Sigma)^{-1} \]

\[ w_{100} = P^{-1} Q (\delta \Sigma)^{-1} \]

And combining some of the terms

\[ \hat{w} = \Pi (\delta \Sigma)^{-1} + \left[ \frac{1}{(1+\alpha)} \right] \left[ P^{-1} Q (\delta \Sigma)^{-1} - \Pi (\delta \Sigma)^{-1} \right] \]

Substituting in \( w_{mk} \) and \( w_{100} \).

\[ \hat{w} = w_{mk} + \left[ \frac{1}{(1+\alpha)} \right] \left[ w_{100} - w_{mk} \right] \]

And looking at formula (39)

\[ \text{confidence} = (\hat{w} - w_{mk}) / (w_{100} - w_{mk}) \]

We see the

\[ \text{confidence} = \frac{1}{(1+\alpha)} \]

And if we solve for \( \alpha \)

\[ \alpha = (1 - \text{confidence}) / \text{confidence} \]

To check the results for each view, we then solve for the posterior estimated returns using formula (27) and then plug them back into formula (39).

The result of the process will be values of \( \omega_i \) for each individual view. All the values can be rolled up into a single \( \Omega \) matrix and the Black-Litterman process resumed. This greatly simplifies the process of specifying the views, and also removes any dependency of \( \tau \) on the returns. Idzorek carries \( \tau \) along, but his method is insensitive to the value of \( \tau \) used.

**An Example of Idzorek's Extension**

Idzorek describes the steps required to implement his extension in his paper, but does not provide a worked example. In this section I will work through his example in detail from where he leaves off in his paper.

© Copyright 2008, Jay Walters
Idzorek's example includes 3 views:

- International Dev Equity will have absolute excess return of 5.25%, Confidence 25.0%
- International Bonds will outperform US bonds by 25bps, Confidence 50.0%
- US Growth Equity will outperform US Value Equity by 2%, Confidence 65.0%

In order to use Idzorek's method we will need to operate on each of these views by itself. We will use the equilibrium weights as a base, then we will apply each view in turn with 100% certainty (Ω = 0) and compute new weights. We will assume a linear model for uncertainty, and stipulate that the change in posterior mean for a x% certainty in the view will be x% of the difference between the equilibrium weight and the 100% certainty weight.

Tables 4, 5 and 6, below, shows results for each view. For each view the first table (A) shows the results given the standard value for Ω which is diag(P(τΣ)P'). This gives an implied confidence level using Idzorek's linear model of confidence to weights. Note for view 1, the implied confidence in the A table is not close to the desired level. The second table (B) shows the results after calibrating the Ω value for the desired confidence level.

Table 4A – Results for View 1

<table>
<thead>
<tr>
<th>Asset</th>
<th>ω1</th>
<th>W_mkt</th>
<th>w*</th>
<th>W_100%</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intl Dev Equity</td>
<td>0.000708875</td>
<td>24.18%</td>
<td>27.77%</td>
<td>29.28%</td>
<td>70.47%</td>
</tr>
</tbody>
</table>

Table 4B – Calibrated Results for View 1

<table>
<thead>
<tr>
<th>Asset</th>
<th>ω1</th>
<th>W_mkt</th>
<th>w*</th>
<th>W_100%</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intl Dev Equity</td>
<td>0.002124023</td>
<td>24.18%</td>
<td>25.46%</td>
<td>29.28%</td>
<td>25.02%</td>
</tr>
</tbody>
</table>

Table 5A – Results for View 2

<table>
<thead>
<tr>
<th>Asset</th>
<th>ω2</th>
<th>W_mkt</th>
<th>w*</th>
<th>W_100%</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Bonds</td>
<td>0.000140650</td>
<td>19.34%</td>
<td>29.89%</td>
<td>38.78%</td>
<td>54.26%</td>
</tr>
<tr>
<td>Intl Bonds</td>
<td>0.000140650</td>
<td>26.13%</td>
<td>15.58%</td>
<td>6.69%</td>
<td>54.26%</td>
</tr>
</tbody>
</table>

Table 5B – Calibrated Results for View 2

<table>
<thead>
<tr>
<th>Asset</th>
<th>ω2</th>
<th>W_mkt</th>
<th>w*</th>
<th>W_100%</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Bonds</td>
<td>0.000140667</td>
<td>19.34%</td>
<td>29.06%</td>
<td>38.78%</td>
<td>50.00%</td>
</tr>
<tr>
<td>Intl Bonds</td>
<td>0.000140667</td>
<td>26.13%</td>
<td>16.41%</td>
<td>6.69%</td>
<td>50.00%</td>
</tr>
</tbody>
</table>
### Table 6A - Results for View 3

<table>
<thead>
<tr>
<th>Asset</th>
<th>$\omega_3$</th>
<th>$w_{mkt}$</th>
<th>$w^*$</th>
<th>$w_{100%}$</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>US LG</td>
<td>.000865628</td>
<td>12.09%</td>
<td>9.37%</td>
<td>8.09%</td>
<td>68.06%</td>
</tr>
<tr>
<td>US LV</td>
<td>.000865628</td>
<td>12.09%</td>
<td>14.81%</td>
<td>16.09%</td>
<td>68.06%</td>
</tr>
<tr>
<td>US SG</td>
<td>.000865628</td>
<td>1.34%</td>
<td>1.04%</td>
<td>.90%</td>
<td>68.06%</td>
</tr>
<tr>
<td>US SV</td>
<td>.000865628</td>
<td>1.34%</td>
<td>1.64%</td>
<td>1.78%</td>
<td>68.06%</td>
</tr>
</tbody>
</table>

### Table 6B – Calibrated Results for View 3

<table>
<thead>
<tr>
<th>Asset</th>
<th>$\omega_3$</th>
<th>$w_{mkt}$</th>
<th>$w^*$</th>
<th>$w_{100%}$</th>
<th>Implied Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>US LG</td>
<td>.000466003</td>
<td>12.09%</td>
<td>9.49%</td>
<td>8.09%</td>
<td>65.01%</td>
</tr>
<tr>
<td>US LV</td>
<td>.000466003</td>
<td>12.09%</td>
<td>14.69%</td>
<td>16.09%</td>
<td>65.01%</td>
</tr>
<tr>
<td>US SG</td>
<td>.000466003</td>
<td>1.34%</td>
<td>1.05%</td>
<td>.90%</td>
<td>65.01%</td>
</tr>
<tr>
<td>US SV</td>
<td>.000466003</td>
<td>1.34%</td>
<td>1.63%</td>
<td>1.78%</td>
<td>65.01%</td>
</tr>
</tbody>
</table>

Then we use the freshly computed values for the $\Omega$ matrix with all views specified together and arrive at the following final result shown in Table 7 blending all 3 views together.
<table>
<thead>
<tr>
<th>Asset</th>
<th>View 1</th>
<th>View 2</th>
<th>View 3</th>
<th>μ</th>
<th>σ</th>
<th>(w_{mkt})</th>
<th>Posterior Weight</th>
<th>change</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Bonds</td>
<td>0.0</td>
<td>-1.0</td>
<td>0.0</td>
<td>.1</td>
<td>3.2</td>
<td>19.3%</td>
<td>29.6%</td>
<td>10.3%</td>
</tr>
<tr>
<td>Intl Bonds</td>
<td>0.0</td>
<td>1.0</td>
<td>0.0</td>
<td>.5</td>
<td>8.5</td>
<td>26.1%</td>
<td>15.8%</td>
<td>10.3%</td>
</tr>
<tr>
<td>US LG</td>
<td>0.0</td>
<td>0.0</td>
<td>0.9</td>
<td>6.3</td>
<td>24.5</td>
<td>12.1%</td>
<td>8.9%</td>
<td>3.2%</td>
</tr>
<tr>
<td>US LV</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.9</td>
<td>4.2</td>
<td>17.2</td>
<td>12.1%</td>
<td>15.2%</td>
<td>3.2%</td>
</tr>
<tr>
<td>US SG</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>7.3</td>
<td>32.0</td>
<td>1.3%</td>
<td>1.0%</td>
<td>-.4%</td>
</tr>
<tr>
<td>US SV</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.1</td>
<td>3.8</td>
<td>17.9</td>
<td>1.3%</td>
<td>1.7%</td>
<td>.4%</td>
</tr>
<tr>
<td>Intl Dev</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>4.8</td>
<td>16.8</td>
<td>24.2%</td>
<td>26.0%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Intl Emg</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>6.6</td>
<td>28.3</td>
<td>3.5%</td>
<td>3.5%</td>
<td>-.0%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>101.8%</td>
<td></td>
</tr>
<tr>
<td>Return</td>
<td>5.2</td>
<td>.2</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omega/tau</td>
<td>.08496</td>
<td>.00563</td>
<td>.01864</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Fusai and Meucci's Method of Measuring Consistency

This section of the document will discuss reproducing the results of Fusai and Meucci (2003). In their paper they present a way to quantify the statistical difference between the posterior return estimates and the prior estimates. This provides a way to calibrate the uncertainty of the views and ensure that the posterior estimates are not extreme when viewed in the context of the prior equilibrium estimates.

They propose the use of the Mahalanobis distance which is the multi-dimensional analog of the z-score (I include \(\tau\) to match the Black-Litterman reference model).

\[
M(q) = (E(r) - \Pi)(\tau\Sigma)^{-1}(E(r) - \Pi)
\]

The Mahalanobis distance is distributed as a chi-square distribution with \(n\) degrees of freedom (\(n\) is the number of assets). It is essentially measuring the distance from the prior, \(\Pi\), to the estimated returns, \(E(r)\), normalized by the uncertainty in the estimate. We use the covariance matrix computed from historical data as the uncertainty.

Thus the probability of this event occurring can be computed as:

\[
P(q) = 1 - F(M(q))
\]
Where $F(M(q))$ is the CDF of the chi square distribution of $M(q)$ with $n$ degrees of freedom.

Finally, in order to identify which views contribute most highly to the distance away from the equilibrium, we can also compute sensitivities of the probability to each view. We use the chain rule to compute the partial derivatives

$$
\frac{\partial P(q)}{\partial q} = \frac{\partial P}{\partial M} \frac{\partial M}{\partial \mu_{BL}} \frac{\partial \mu_{BL}}{\partial q}
$$

(42) $$
\frac{\partial P(q)}{\partial q} = f(M)[-2(\mu_{BL} - \mu)](P(\tau \Sigma)^{-1}P')P]
$$

Where $F(M)$ is the PDF of the chi square distribution with $n$ degrees of freedom for $M(q)$.

Fusai and Meucci (2003) make two modifications to the Black-Litterman model, first they use the alternative reference model and second they set $\Omega = \alpha P \Sigma P^T$ where $\alpha$ is a positive scalar. This is in contrast to other variants of the Black-Litterman model that make $\Omega$ diagonal in order to make it simpler to estimate. By keeping $\Omega$ exactly proportional to $\Sigma$ they don't lose the stability and they can use some reduced forms of the formulas, such to formula (36). By using the alternative reference model they also fix the parameter $\tau$ at 1, which means that $\Sigma$ is directly the variance of the estimate. In their paper they use versions of formulas (40) and (42) without $\tau$ because of their use of the alternative reference model. Independent of their assumptions, their technique of computing the probability for the views, and the sensitivities of the probability to the individual views is useful for all methods of implementing the Black-Litterman model.

They work an example in their paper which results in an initial probability of 94% that the posterior is consistent with the prior. They specify that their investor desires this probability to be no less than 95%, and thus they would adjust their views to bring the probability in line. Given that they also compute sensitivities, their investor can identify which views are sensitive and in which direction (diverging or converging with the prior).

We can apply this method to any Black-Litterman problem. If we apply their method to the example from He and Litterman (1999) we see that the probability is greater than 99% that these views are consistent with the prior. This is because the views as measured by their impact on the posterior return estimates are not extreme. The views do have a significant impact on the optimal weights from the mean-variance optimizer, but this is to be expected given how sensitive mean-variance optimization is to small changes in estimated returns.

**Two-Factor Black-Litterman**

Krishnan and Mains (2005) developed an extension to the alternate reference model which allows the incorporation of additional uncorrelated market factors. The main point they make is that the Black-Litterman model measures risk, like all MVO approaches, as the covariance of the assets. They advocate for a richer measure of risk. They specifically focus on a recession indicator, given the thesis that many investors want assets which perform well during recessions and thus there is a positive risk premium associated with holding assets which do poorly during recessions. Their approach is general and can be applied to one or more additional market factors given that the market has zero beta to the factor and the factor has a non-zero risk premium.

They start from the standard quadratic utility function (5), but add an additional term for the new market factor(s).
(43) \[ U = w^T \Pi - \left( \frac{\delta_0}{2} \right) w^T \Sigma w - \sum_{j=1}^{n} \delta_j w^T \beta_j \]

\( U \) is the investors utility, this is the objective function during portfolio optimization.
\( w \) is the vector of weights invested in each asset
\( \Pi \) is the vector of equilibrium excess returns for each asset
\( \Sigma \) is the covariance matrix for the assets
\( \delta_0 \) is the risk aversion parameter of the market
\( \delta_j \) is the risk aversion parameter for the j-th additional risk factor
\( \beta_j \) is the vector of exposures to the j-th additional risk factor

Given their utility function as shown in formula (43) we can take the first derivative with respect to \( w \) in order to solve for the equilibrium asset returns.

(44) \[ \Pi = \delta_0 \Sigma w + \sum_{j=1}^{n} \delta_j \beta_j \]

Comparing this to formula (6), the simple reverse optimization formula, we see that the equilibrium excess return vector \( \Pi \) is a linear composition of (6) and a term linear in the \( \beta_j \) values. This matches our intuition as we expect assets exposed to this extra factor to have additional return above the equilibrium return.

We will further define the following quantities:
- \( r_m \) as the return of the market portfolio.
- \( f_j \) as the time series of returns for the factor
- \( r_j \) as the return of the replicating portfolio for risk factor \( j \).

In order to compute the values of \( \delta \) we will need to perform a little more algebra. Given that the market has no exposure to the factor, then we can find a weight vector, \( v_0 \), such that \( v_0^T \beta_j = 0 \). In order to find \( v_j \) we perform a least squares fit of \( \| f_j - v_j^T \Pi \| \) subject to the above constraint. \( v_0 \) will be the market portfolio, and \( v_0 \beta_j = 0 \) \( \forall \) \( j \) by construction. We can solve for the various values of \( \delta \) by multiplying formula (44) by \( v \) and solving for \( \delta_0 \).

\[ v_0^T \Pi = \delta_0 v_0^T \Sigma v_0 + \sum_{j=1}^{n} \delta_j v_0^T \beta_j \]

By construction \( v_0 \beta_j = 0 \), and \( v_0 \Pi = r_m \), so
\[ \delta_0 = \frac{r_m}{v_0^T \Sigma v_0} \]

For any \( j \geq 1 \) we can multiply formula (44) by \( v_j \) and substitute \( \delta_0 \) to get
\[ v_j^T \Pi = \delta_0 v_j^T \Sigma v_j + \sum_{i=1}^{n} \delta_i v_j^T \beta_i \]

Because these factors must all be independent and uncorrelated, then \( v_i \beta_j = 0 \) \( \forall \) \( i \neq j \) so we can solve for each \( \delta_i \).
\[ \delta_j = \frac{(r_j - \delta_0 v_j^T \Sigma v_j)}{(v_j^T \beta_j)} \]

The authors raise the point that this is only an approximation because the quantity \( \| f_j - v_j^T \Pi \| \) may not be identical to 0. The assertion that \( v_i \beta_j = 0 \ \forall \ i \neq j \) may also not be satisfied for all \( i \) and \( j \). For the case of a single additional factor, we can ignore the latter issue.

In order to transform these formulas so we can directly use the Black-Litterman model, Krishnan and Mains change variables, letting

\[ \hat{\Pi} = \Pi - \sum_{j=1}^{n} \delta_j \beta_j \]

Substituting back into (43) we are back to the standard utility function

\[ U = w^T \hat{\Pi} - \frac{\delta_0}{2} w^T \Sigma w \]

and from formula (10)

\[ P \hat{\Pi} = P (\Pi - \sum_{j=1}^{n} \delta_j \beta_j) \]

\[ P \hat{\Pi} = P \Pi - \sum_{j=1}^{n} \delta_j P \beta_j \]

thus

\[ \hat{Q} = Q - \sum_{j=1}^{n} \delta_j P \beta_j \]

We can directly substitute \( \hat{\Pi} \) and \( \hat{Q} \) into formula (27) for the posterior returns in the Black-Litterman model in order to compute returns given the additional factors. Note that these additional factor(s) do not impact the posterior variance in any way.

Krishnan and Mains work an example of their model for world equity models with an additional recession factor. This factor is comprised of the Altman Distressed Debt index and a short position in the S&P 500 index to ensure the market has a zero beta to the factor. They work through the problem for the case of 100% certainty in the views. They provide all of the data needed to reproduce their results given the set of formulas in this section. In order to perform all the regressions, one would need to have access to the Altman Distressed Debt index along with the other indices used in their paper.

**Future Directions**

Future directions for this research include reproducing the results from the original papers, either Black and Litterman (1991) or Black and Litterman (1992). These results have the additional complication of including currency returns and partial hedging.

Later versions of this document should include more information on process and a synthesized model containing the best elements from the various authors. A full example from the CAPM equilibrium, through the views to the final optimized weights would be useful, and a worked example of the two factor model from Krishnan and Mains (2005) would also be useful.

The article, Meucci (2006) on the use of non-normal views and Black-Litterman provides a nice
extension to Black-Litterman for non-normal views such as one might find in the alternative investment or derivatives world. One drawback to his approach is the use of numerical methods rather than analytical methods. Gaining a better understanding of this paper would provide useful information. I hope to include a MATLAB/SciLab compatible implementation of the Black-Litterman model at some point in the future, to augment the Java implementation which I currently have.

**Literature Survey**

This section will provide a quick overview of the references to Black-Litterman in the literature. The initial paper, Black and Litterman (1991) provides some discussion of the model, but does not include significant details and also does not include all the data necessary to reproduce their results. They introduce a parameter, weight on views, which is used in a few of the other papers but not clearly defined. It appears to be the fraction \( \frac{P^T \Omega^{-1} P}{(\tau \Sigma)^{-1} + P^T \Omega^{-1} P} \). This represents the weight of the view returns in the mixing. As \( \Omega \to 0 \), then the weight on views \( \to 100\% \).

Their second paper on the model, Black and Litterman (1992), provides a good discussion of the model along with the main assumptions. The authors present several results and most of the input data required to generate the results, however they do not document all their assumptions in any easy to use fashion. As a result, it is not trivial to reproduce their results. They provide some of the key equations required to implement the Black-Litterman model, but they do not provide any equations for the posterior variance.

He and Litterman (1999) provide a clear and reproducible discussion of Black-Litterman. There are still a few fuzzy details in their paper, but along with Idzorek (2004) one can recreate the mechanics of the Black-Litterman model. Using the He and Litterman source data, and their assumptions as documented in their paper one can reproduce their results.

Idzorek (2004) provides his inputs and assumptions allowing his results to be reproduced. During this process of reproducing their results, I identified the fact that Idzorek does not handle the posterior variance the same way as He and Litterman.

Bevan and Winkelmann (1998) and the chapter from Litterman's book Litterman, et al, (2003) do not shed any further light on the details of the algorithm. Neither provides the details required to build the model or to reproduce any results they might discuss. Bevan and Winkelmann (1998) provide details on how they use Black-Litterman as part of their broader Asset Allocation process at Goldman Sachs, including some calibrations of the model which they perform. This is useful information for anybody planning on building Black-Litterman into an ongoing asset allocation process.

Satchell and Scowcroft (2000) claim to demystify Black-Litterman, but they don't provide enough details to reproduce their results, and they seem to have a very different view on the parameter \( \tau \) than the other authors do. I see no intuitive reason to back up their assertion that \( \tau \) should be set to 1. They provide a detailed derivation of the Black-Litterman 'master formula'..

Christadoulakis (2002) and Da and Jagannathan (2005) are teaching notes for Asset Allocation classes. Christadoulakis (2002) provides some details on the Bayesian mechanisms, the assumptions of the model and enumerates the key formulas for posterior returns. Da and Jagannathan (2005) provides some discussion of an excel spreadsheet they build and works through a simple example.

Herold (2003) provides an alternative view of the problem where he examines optimizing alpha generation, essentially specifying that the sample distribution has zero mean.. He provides some
additional measures which can be used to validate that the views are reasonable.

Koch (2005) is a powerpoint presentation on the Black-Litterman model. It includes derivations of the 'master formula' and the alternative form under 100% certainty. He does not mention posterior variance, or show the alternative form of the 'master formula' under uncertainty (general case).

Krishnan and Mains (2005) provide an extension to the Black-Litterman model for an additional factor which is uncorrelated with the market. They call this the Two-Factor Black-Litterman model and they show an example of extending Black-Litterman with a recession factor. They show how it intuitively impacts the expected returns computed from the model.

Mankert (2006) provides a nice solid walk through of the model and provides a detailed transformation between the two specifications of the Black-Litterman 'master formula' for the estimated asset returns. She also provides some new intuition about the value τ, from the point of view of sampling theory.

Meucci (2006) provides a method to use non-normal views in Black-Litterman. I have not had the time to dig into this paper and understand exactly what he does. He does have MATLAB code for his example on his website with the paper.

Several of the other authors refer to a reference Firoozy and Blamont, Asset Allocation Model, Global Markets Research, Deutsche Bank, July 2003. I have been unable to find a copy of this document. I will at times still refer to this document based on comments by other authors. After reading other authors references to their paper, I believe my approach to the problem is somewhat similar to theirs.

References

Many of these references are available on the Internet. I have placed a Black-Litterman resources page on my website, (www.blacklitterman.org) with links to many of these papers.


© Copyright 2008, Jay Walters


Appendix A

This appendix includes the derivation of the Black-Litterman master formula using Theil's Mixed Estimation approach which is based on Generalized Least Squares.

Theil's Mixed Estimation Approach

This approach is from [Theil and Goldberger, 1963] and is similar to the reference in the original [Black and Litterman, 1992] paper. [Koch, 2005] also includes a derivation similar to this.

If we start with a prior distribution for the returns. Assume a linear model such as

A.1 \[ x = E(r) + v \]

Where \( x \) is the mean of the prior return distribution, \( E(r) \) is the expected return and \( u \) is the normally distributed residual with mean 0 and variance \((S/n)\).

Next we consider some additional information, the conditional distribution.

A.2 \[ \mu = E(r) + v \]

Where \( \mu \) is the mean of the conditional distribution and \( v \) is the normally distributed residual with mean 0 and variance \( \Sigma \).

Both \( S \) and \( \Sigma \) are assumed to be non-singular.

We can combine the prior and conditional information by writing:

A.3 \[ \begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} I \\ I \end{bmatrix} E(r) + \begin{bmatrix} u \\ v \end{bmatrix} \]

Where the expected value of the residual is 0, and the expect value of the variance is

\[ E\left( \begin{bmatrix} u \\ v \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix} \right) = \begin{bmatrix} S/n & 0 \\ 0 & \Sigma \end{bmatrix} \]

We can then apply the generalized least squares procedure, which leads to estimating \( E(r) \) as

A.4 \[ \hat{E}(r) = \begin{bmatrix} I & I \end{bmatrix} \begin{bmatrix} S/n & 0 \\ 0 & \Sigma \end{bmatrix}^{-1} \begin{bmatrix} I' \\ I' \end{bmatrix} \begin{bmatrix} S/n & 0 \\ 0 & \Sigma \end{bmatrix}^{-1} \begin{bmatrix} x \\ \mu \end{bmatrix} \]

This can be rewritten without the matrix notation as

A.5 \[ \hat{E}(r) = \left( (S/n)^{-1} + \Sigma^{-1} \right)^{-1} \left[ x(S/n)^{-1} + \mu \Sigma^{-1} \right] \]

We can derive the expression for the variance using similar logic. Given that the variance is the expectation of \( (\hat{E}(r) - E(r))^2 \), then we can start by substituting formula A.3 into A.5

A.6 \[ \hat{E}(r) = \left( (S/n)^{-1} + \Sigma^{-1} \right)^{-1} \left[ (E(r)+u)(S/n)^{-1} + (E(r)+v) \Sigma^{-1} \right] \]

This simplifies to

\[ \hat{E}(r) = \left( (S/n)^{-1} + \Sigma^{-1} \right)^{-1} \left[ E(r)(S/n)^{-1} + E(r) \Sigma^{-1} \right] + \left( (S/n)^{-1} + \Sigma^{-1} \right)^{-1} \left[ u(S/n)^{-1} + v \Sigma^{-1} \right] \]

© Copyright 2008, Jay Walters
\[ \hat{E}(r) = E(r)[(Sln)^{-1} + \Sigma^{-1}]^{-1}[(Sln)^{-1} + \Sigma^{-1}] + [(Sln)^{-1} + \Sigma^{-1}]^{-1}[u(Sln)^{-1} + v \Sigma^{-1}] \]

\[ \hat{E}(r) = E(r) + [(Sln)^{-1} + \Sigma^{-1}]^{-1}[u(Sln)^{-1} + v \Sigma^{-1}] \]

**A.7**

\[ \hat{E}(r) - E(r) = [(Sln)^{-1} + \Sigma^{-1}]^{-1}[u(Sln)^{-1} + v \Sigma^{-1}] \]

The variance is the expectation of formula A.7 squared.

\[ E(\hat{E}(r) - E(r))^2 = [(Sln)^{-1} + \Sigma^{-1}]^{-1}[u(Sln)^{-1} + v \Sigma^{-1}]^2 \]

\[ E(\hat{E}(r) - E(r))^2 = [(Sln)^{-1} + \Sigma^{-1}]^2[u^2(Sln)^{-2} + v^2 \Sigma^{-2} + uv(Sln)^{-1} \Sigma^{-1}] \]

We know from above that \( E(\hat{u}\hat{u}') = (Sln) \), \( E(\hat{v}\hat{v}') = \Sigma \) and \( E(\hat{u}\hat{v}') = 0 \) because \( u \) and \( v \) are independent variables, so substituting

\[ E(\hat{E}(r) - E(r))^2 = [(Sln)^{-1} + \Sigma^{-1}]^2[(Sln)(Sln)^{-2} + \Sigma \Sigma^{-2} + 0] \]

**A.8**

\[ E(\hat{E}(r) - E(r))^2 = [(Sln)^{-1} + \Sigma^{-1}]^{-1} \]
Appendix B

This appendix contains a derivation of the Black-Litterman master formula using the standard Bayesian approach for modeling the posterior of two normal distributions. One additional derivation is in [Mankert, 2006] where she derives the Black-Litterman 'master formula' from Sampling theory, and also shows the detailed transformation between the two forms of this formula.

The PDF Based Approach

The PDF Based Approach follows a Bayesian approach to computing the PDF of the posterior distribution, when the prior and conditional distributions are both normal distributions. This section is based on the proof shown in [DeGroot, 1970]. This is similar to the approach taken in [Satchell and Scowcroft, 2000].

The method of this proof is to examine all the terms in the PDF of each distribution which depend on $E(r)$, neglecting the other terms as they have no dependence on $E(r)$ and thus are constant with respect to $E(r)$.

Starting with our prior distribution, we derive an expression proportional to the value of the PDF.

$P(A) \propto N(x,S/n)$ with $n$ samples from the population.

So $\xi(x)$ the PDF of $P(A)$ satisfies

$$B.1 \quad \xi(x) \propto \exp \left( \frac{S}{n} - 1 \left( E(r) - x \right)^2 \right)$$

Next, we consider the PDF for the conditional distribution.

$P(B|A) \propto N(\mu, \Sigma)$

So $\xi(\mu|x)$ the PDF of $P(B|A)$ satisfies

$$B.2 \quad \xi(\mu|x) \propto \exp \left( \Sigma^{-1} \left( E(r) - \mu \right)^2 \right)$$

Substituting B.1 and B.2 into formula (1) from the text, we have an expression which the PDF of the posterior distribution will satisfy.

$$B.3 \quad \xi(x|\mu) \propto \exp \left( - \left( \Sigma^{-1} \left( E(r) - \mu \right)^2 + \frac{S}{n} \right)^{-1} \left( E(r) - x \right)^2 \right) ,$$

or

$$\xi(x|\mu) \propto \exp(-\Phi)$$

Considering only the quantity in the exponent and simplifying

$$\Phi = \left( \Sigma^{-1} \left( E(r) - \mu \right)^2 + \frac{S}{n} \right)^{-1} \left( E(r) - x \right)^2$$

$$\Phi = \Sigma^{-1} \left( E(r)^2 - 2 E(r) \mu + \mu^2 \right) + \frac{S}{n} \left( E(r)^2 - 2 E(r) x + x^2 \right)$$

$$\Phi = E(r)^2 \left( \Sigma^{-1} + \frac{S}{n} \right)^{-1} - 2 E(r) (\mu \Sigma^{-1} + x (S/n)^{-1}) + \Sigma^{-1} \mu^2 + (S/n)^{-1} x^2$$

If we introduce a new term $y$, where
B.4 \[ y = \frac{[\mu \Sigma^{-1} + x(Sln)^{-1}]}{(\Sigma^{-1} + (Sln)^{-1})} \]

and then substitute in the second term

\[ \Phi = E(r)^2 (\Sigma^{-1} + (Sln)^{-1}) - 2E(r)y(\Sigma^{-1} + (Sln)^{-1}) + \Sigma^{-1} \mu^2 + (Sln)^{-1} \]

Then add

\[ 0 = y^2(\Sigma^{-1} + (Sln)^{-1}) - (\mu \Sigma^{-1} + x(Sln)^{-1})^2(\Sigma^{-1} + (Sln)^{-1})^{-1} \]

\[ \Phi = E(r)^2 (\Sigma^{-1} + (Sln)^{-1}) - 2E(r)y(\Sigma^{-1} + (Sln)^{-1}) + \Sigma^{-1} \mu^2 + (Sln)^{-1} \]

\[ + y^2(\Sigma^{-1} + (Sln)^{-1}) - (\mu \Sigma^{-1} + x(Sln)^{-1})^2(\Sigma^{-1} + (Sln)^{-1})^{-1} \]

\[ \Phi = E(r)^2 (\Sigma^{-1} + (Sln)^{-1}) - 2E(r)y(\Sigma^{-1} + (Sln)^{-1}) + y^2(\Sigma^{-1} + (Sln)^{-1}) \]

\[ + \Sigma^{-1} \mu^2 + (Sln)^{-1} x^2 - (\mu \Sigma^{-1} + x(Sln)^{-1})^2(\Sigma^{-1} + (Sln)^{-1})^{-1} \]

The second term has no dependency on E(r), thus it can be included in the proportionality factor and we are left with

B.5 \[ \xi(x|\mu) \propto \exp \left( -[\Sigma^{-1} + (Sln)^{-1}]^{-1} E(R) - y^2 \right) \]

Thus the posterior mean is y as defined in formula A.12, and the variance is

B.6 \[ (\Sigma^{-1} + (Sln)^{-1})^{-1} \]
Appendix C

This appendix provides a derivation of the alternate format of the posterior variance. This format does not require the inversion of Ω, and thus is more stable computationally.

\[ ((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} = ((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} \]

\[ (((τΣ)^{-1} + P^TΩ^{-1}P)^{-1})^T(P^T)^{-1} \]

\[ (((τΣ)^{-1} + P^T)^{-1}(τΣ^P)^{-1} + P^TΩ^{-1}P)^{-1}(P^T)^{-1} \]

\[ (((τΣ^P)^{-1} + Ω^{-1}P)^{-1}(τΣ)(τΣ)^{-1}(P^T)^{-1} \]

\[ (((τΣ)(τΣ)^{-1} + Ω^{-1}P)^{-1}(τΣ)(τΣ)^{-1}((P^T)^{-1} \]

\[ (((τΣ^P)^{-1} + Ω^{-1}P)^{-1}(τΣ)(τΣ)^{-1}(P^T)^{-1} \]

\[ (((τΣ)(τΣ)^{-1} + Ω^{-1}P)^{-1}(τΣ)(P^TτΣ)^{-1} \]

\[ (((τΣ^P)^{-1} + Ω^{-1}P)^{-1}(τΣ)(P^TτΣ)^{-1} \]

\[ (τΣ)(P^TτΣ)^{-1} = ((τΣ^P)^{-1} + Ω^{-1}P)((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} \]

\[ (τΣ)(P^TτΣ)^{-1} - (τΣ^P)^{-1}((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} \]

\[ (τΣ)(P^TτΣ)^{-1} - (τΣ^P)^{-1}((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} \]

\[ (τΣ)(P^TτΣ)^{-1} - (τΣ^P)^{-1}((τΣ)^{-1} + P^TΩ^{-1}P)^{-1} \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]

\[ (τΣ)(P^TτΣ)^{-1} - (P(τΣ)P^T + Ω)^{-1}P(τΣ) \]
Appendix D

This appendix presents a derivation of the alternate formulation of the Black-Litterman master formula for the posterior expected return. Starting from formula (26) we will derive formula (27).

$$E(r) = [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} [(\tau \Sigma)^{-1} \Pi + P^T \Omega^{-1} Q]$$

Separate the parts of the second term

$$E(r) = [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} [P \tau \Sigma] (\tau \Sigma)^{-1} \Pi + [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} (P^T \Omega^{-1} Q)$$

Replace the precision term in the first term with the alternate form

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \tau \Sigma] (\tau \Sigma)^{-1} \Pi + [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + (\tau \Sigma) (\tau \Sigma)^{-1} [(\tau \Sigma)^{-1} + P^T \Omega^{-1} P]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + (\tau \Sigma) [I_n + P^T \Omega^{-1} P \tau \Sigma]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + \tau \Sigma [I_n + P^T \Omega^{-1} P \tau \Sigma]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + \tau \Sigma [P^T \Omega^{-1} P \tau \Sigma]^{-1} (P^T \Omega^{-1} Q)$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + \tau \Sigma [\Omega (P^T)^{-1} + P \tau \Sigma]^{-1} Q$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + \tau \Sigma [P^T (P^T)^{-1} + \Omega (P^T)^{-1} + P \tau \Sigma]^{-1} Q$$

$$E(r) = [P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} P \Pi] + \tau \Sigma [P^T [\Omega + P \tau \Sigma P^T]^{-1} Q]$$

Voila, the alternate form of the Black-Litterman formula for expected return.

$$E(r) = P \tau \Sigma P^T [P \tau \Sigma P^T + \Omega]^{-1} \Pi [Q - P \Pi]$$

© Copyright 2008, Jay Walters
Appendix E

This section of the document summarizes the steps required to implement the Black-Litterman model. You can use this road map to implement either the He and Litterman version of the model, or the Idzorek version of the model. The Idzorek version of the Black-Litterman model leaves out two steps.

Given the following inputs

- \( w \) Equilibrium weights for each asset class. Derived from capitalization weighted CAPM Market portfolio,
- \( \Sigma \) Matrix of covariances between the asset classes. Can be computed from historical data.
- \( r_f \) Risk free rate for base currency
- \( \delta \) The risk aversion coefficient of the market portfolio. This can be assumed, or can be computed if one knows the return and standard deviation of the market portfolio.
- \( \tau \) A measure of uncertainty of the equilibrium variance. Usually set to a small number of the order of 0.025 – 0.050.

First we use reverse optimization to compute the vector of equilibrium returns, \( \Pi \) using formula (6).

\[
\Pi = \delta \Sigma w
\]

Then we formulate the investors views, and specify \( P, \Omega \) and \( Q \). Given \( k \) views and \( n \) assets, then \( P \) is a \( k \times n \) matrix where each row sums to 0 (relative view) or 1 (absolute view). \( Q \) is a \( k \times 1 \) vector of the excess returns for each view. \( \Omega \) is a diagonal \( k \times k \) matrix of the variance of the views, or the confidence in the views. As a starting point, most authors call for the values of \( \omega_i \) to be set equal to \( p^T \tau \Sigma p \) (where \( p \) is the row from \( P \) for the specific view).

Next assuming we are uncertain in all the views, we apply the Black-Litterman 'master formula' to compute the posterior estimate of the returns using formula (27).

\[
E(r) = \Pi + \tau \Sigma P^T \left[ (P \tau \Sigma P^T) + \Omega \right]^{-1} Q - P \Pi
\]

This following two steps are not needed when using the alternative reference model.

Next we must compute the posterior variance using formula (32).

\[
M = \tau \Sigma - \tau \Sigma P^T \left[ P \Sigma P^T + \Omega \right]^{-1} P \tau \Sigma
\]

Closely followed by the computation of the sample variance from formula (29).

\[
\Sigma_p = \Sigma + M
\]

And now we can compute the portfolio weights for the optimal portfolio on the unconstrained efficient frontier from formula (8).

\[
w = \Pi (\delta \Sigma)^{-1} \quad w = \Pi (\delta \Sigma_p)^{-1}
\]