# The Black-Litterman Model: A Detailed Exploration

Revised: January 27, 2008

© Copyright 2008 - Jay Walters, CFA
jwalters@computer.org


## *Abstract*

This paper attempts to demystify the Black-Litterman model through brute force. It shows the Bayesian origins of the model, and works through each piece of the model in detail. Several derivations and transformations of the formulas are shown in the body and appendices. Next, we reverse engineer the implementations used in several of the major papers and discuss the differences therein. The major papers considered are: [Black and Litterman, 1992], [He and Litterman, 1999], [Satchell and Scowcroft, 2000], [Fusai and Meucci, 2003],[Idzorek, 2004] and [Krishnan and Mains, 2005]. Finally, examples are worked for some of the extensions, e.g. Idzorek' approach to view confidence levels. Where possible examples are worked with the original author's data for the illustrative and comparative purposes.


## *Overview*

This document is divided into several sections. The first section provides a quick overview of the Black-Litterman model. The second section provides an introduction to the relevant portion of Bayesian theory, primarily to define terms that will be used in the rest of the document. I have taken the approach of demonstrating much of the Black-Litterman machinery as natural results from Bayesian theory. This somewhat simplifies the formulas and makes the mathematics a little more intuitive. In this section and several appendices I show the derivations of the various Bayesian results upon which Black-Litterman is is built. References to Black-Litterman specific derivations in the literature are provided. The third section walks the reader through the Black-Litterman model and it's various formulas. The fourth section provides a discussion of the results from the various papers, and how to reproduce them. The fifth section deals with extensions to the model, and includes a worked example of the Idzorek extensions for setting the variance of the views. The final section discusses future directions for this research. The paper concludes with an annotated bibliography.

## *Black-Litterman the Model*

The Black-Litterman model was first published by Fischer Black and Robert Litterman of Goldman Sachs in the Journal of Fixed Income in 1991. A longer and richer paper was then published in 1992 in the Financial Analysts Journal (FAJ). The latter article was then republished by FAJ later in the 1990s. Copies of the second article are widely available on the Internet. It shows some of the derivation of the formulas, but does not show all the formulas required. It also includes a rather complex worked example, which is very difficult to reproduce.

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 the application of Bayesian techniques to estimate returns. Before Black-Litterman, Bayesian work on return estimation started either with an 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, it provides a clear way to specify investors views and to blend the investors views with prior information using Bayesian techniques. This process estimates expected 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. 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 in order to compute a covariance matrix. 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. Finally, the investor needs to quantify their views so that they can be applied and new return estimates computed. Given the limited availability of market capitalization data for illiquid asset classes, e.g. real estate, private equity, it is for most investors that is the most difficult piece of data to gather. 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.

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. [Herold, 2004] provides insights into how implied returns can be computed in the presence of simple constraints such as full investment ($\Sigma w = 1$) and no short selling ($w \geq 0$). The argument can also be made that because the Black-Litterman model relies on excess returns, no budget constraint is required as any excess (shortfall) is long (short) the risk free asset which has expected return 0 and covariance 0 with all assets. This is also consistent with a Bayesian investor who may not wish to be 100% invested in the market due to uncertainty about the results.

For the ensuing discussion, we will consider that the investors views are the conditional distribution, and the CAPM equilibrium distribution is the prior 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. [Satchell and Scowcroft, 2000] also work the problem from this point of view (opposite of [Black and Litterman, 1992]). In the end this difference is not critical as the Bayesian machinery we will develop works for blending any two normal distributions.

 

## *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 for the rest of the paper.

Bayes theory states

(1) $$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

| | |
|---|---|
| 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 know 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 straight forward.

Another core assumption of the Black-Litterman model is that the variance of the prior and the conditional distributions about the true mean are known, but the actual true mean returns are not known. This case, known as “Unknown Mean and Known Variance” is well documented in the Bayesian literature.

We define the significant distributions below:

The prior distribution

(2) $P(A) \sim N(x, S/n)$

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

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

$\Sigma$ 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

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

The variance term in (4) is the variance of the estimated mean about the true 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 precision. Further, the posterior precision is the sum of the prior and conditional precision. Formula (4) 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.

There are two common approaches to the derivation of the posterior distribution. The first is the use of Generalized Least Squares as described in [Theil and Goldberger, 1963]. [Black and Litterman, 1992] refers to this model as Theil's Mixed Estimation model. [Koch, 2005] provides a similar derivation. Other authors, [Satchell and Scowcroft, 2000] for example use a PDF based approach as is shown in their paper, and in [DeGroot, 1970]. Appendix A contains both derivations of formula (4).

If we examine formula (4) in the absence of a conditional distribution, it should collapse into the prior distribution.

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

(5) $$\sim N(x,S/n)$$

As we can see in formula (5), it does indeed collapse to the prior distribution. Another important scenario is the case of 100% certainty of the conditional distribution, where Σ, or some portion of it is 0, and thus Σ is not invertible. We can transform the returns and variance from formula (4) into a form which is more easy to work with in the 100% certainty case.

(6) $$P(A|B) \sim N(x + (S/n)[\Sigma + S/n]^{-1}[\mu - x],[(S/n) - (S/n)(\Sigma + S/n)^{-1}(S/n)])$$

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 Σ is 0 (100% confidence in the views) then the posterior variance will be 0. If Σ is positive infinity (the confidence in the views is 0%) then the posterior variance will be (S/n).

We will revisit equations (4) and (6) later in this paper where we transform these basic equations into the various parts of the Black-Litterman model. Appendices B and C contain derivations of the alternate Black-Litterman formulas from the standard form, analogous to the transformation from (4) to (6).

The final piece of the puzzle is the variance of the posterior returns. Remember that the formulas are computing the variance of the posterior mean about the true mean, they are not computing a posterior estimate of the variance of the asset returns. This variance is the square of the standard error of the posterior distribution. Formula (5) shows that the variance of the posterior distribution in the absence of a conditional distribution is the variance of the prior distribution, S/n.

Given that the error in the estimate of the mean return is independent of the variance of the returns about the true mean, then we could add the two together to get the total variance of returns about the estimated means.

[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

(7) $$M_p = S + M$$

where M is the variance of the posterior distribution about the true mean and S is the known variance of returns.

Or in the absence of views, it simplifies to

$M_p = ((n+1)/n)S$

Later we will see how the Black-Litterman model parameterizes this result.

## *Computing the CAPM Equilibrium Returns*

The process of computing the CAPM equilibrium excess returns is straight forward. These returns will provide the prior distribution for the Black-Litterman model.

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

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

Where

$r_f$ The risk free rate.

$r_m$ The excess return of the market portfolio.

$\beta$ A regression coefficient computed as $\beta = \rho \frac{\sigma_p}{\sigma_m}$

$\alpha$ The residual, or asset specific (idiosyncratic) excess return.

Under the CAPM theory the investor is rewarded for the systemic risk measured by β, but is not rewarded for taking idiosyncratic risk associated with α. This is because within a diversified portfolio the total α should tend to 0 in the limit.

The CAPM theory states that all investors should hold the market portfolio as their risky asset. Depending on their risk aversion they may hold arbitrary fractions of their wealth in the risky asset, and the remainder in the risk-free asset. The 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. Usually we are not concerned with the fact that an implicit budget constraint may be embedded in the model. Because we are using excess returns, any over (under) is long (short) the risk free asset with excess return 0 and covariance with any asset of 0. Later on we will see how a Bayesian investor will divide their investments between the market portfolio and the risk free asset.

We will constrain the problem by asserting that the covariance matrix of the returns, Σ, is known. In practice, this covariance matrix is computed from historical return data. 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.

From this point onwards, 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

* The Sharpe Ratio is the excess return divided by the excess risk, or $(E(r) - r_f) / \sigma$.

notation used in the section on Bayesian theory.

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

(9) $U = w^t\Pi - (\delta/2)w^t\Sigma w$

> 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
> $\delta$ is the risk aversion parameter of the market
> $\Sigma$ 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 (9) with respect to the weights (w) and setting it to 0.

$$dU/dw = \Pi - \delta\Sigma w = 0$$

Solving this for $\Pi$ (the vector of excess returns) yields:

(10) $\Pi = \delta\Sigma w$

In order to use formula (10) we need to have a value for $\delta$ , the risk aversion coefficient of the market. Most of the authors specify the value of $\delta$ 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. We can find $\delta$ by multiplying both sides of (10) by $w^T$ and replacing vector terms with scalar terms.

$$(E(r) - r_f) = \delta\sigma^2$$

(11) $\delta = (E(r) - r_f) / \sigma^2$

> $E(r)$ is the total return on the market portfolio ($E(r) = \Pi + r_f$)
> $r_f$ 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 (11); $E(r)$, $r_f$, and $\sigma^{2.}$ in order to calculate a value for $\delta$. Once we have a value for $\delta$, then we plug w, $\delta$ and $\Sigma$ into formula (10) and generate the set of equilibrium asset returns. Formula (10) 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 (10) to yield the formula for the closed form calculation of the optimal portfolio weights in the absence of constraints.

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

If we feed $\Pi$, $\delta$, and $\Sigma$ back into the above formula, we can solve for the weights (w). If we use historical excess returns rather than equilibrium excess returns, the results will be very sensitive to changes in $\Pi$. With the Black-Litterman model, the weight vector is less sensitive to the reverse optimized $\Pi$ vector. This stability of the optimization process, is one of the strengths of the Black-Litterman model.

We will substitute these values back into formula (2) to arrive at the Black-Litterman prior distribution, but we must still identify what will take the place of n, the number of samples. If we create a new parameter, $\tau$, which represents 1/n in the classical Bayesian model, but for our model represents our confidence in the prior distribution, then we can substitute into formula (2) as follows

 

(12) $P(A) \sim N(\Pi, \tau\Sigma)$

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

## Specifying the Views

This section will describe the process of specifying the investors views on estimated returns. We define the combination of the investors views as the conditional distribution, $P(B|A)$. 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 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.
- 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 $\omega_i$.

Different authors compute the various weights within the view differently, Idzorek likes to use a Capitalization weighed scheme, whereas others use an equal weighted scheme.

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 $\omega > 0$.

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:

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

Given these matrices 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

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

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 (13).

$\Omega$, 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.

There are three main ways to calculate $\Omega$. First, actually compute the variance of the view. This is most easily done by defining a confidence interval around the return, e.g. Asset 2 will have a 3% return with the expectation that 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 $(0.005)^2$.

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.

Second, we can just assume that the variance of the view will be proportional to the variance of the asset returns, just as the variance of the prior distribution is. In this method taken from [He and Litterman, 1999] one just uses the variance of the view computed from the prior distribution,

(14) $\omega_i = p(\tau\Sigma)p^T$

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. By including $\tau$ in the expression, the final solution becomes less dependent on the specific value of $\tau$ selected as well. Several authors have specified the confidence matrix as $\tau\Omega$ in order to manage this interaction.

[Meucci, 2006] specifies $\Omega = \alpha P\tau\Sigma P^t$, where $\alpha$ is an integer greater than or equal to 1. He does not include $\tau$ in his formula because he sets it to 1. He does not require that $\Omega$ be diagonal, but expects that the investors uncertainty in their views is equal or greater to the uncertainty in the equilibrium returns. This formulation simplifies the formulas for posterior mean and variance, because the precisions are now $(\tau\Sigma)^{-1}$ and $(\alpha)^{-1}(\tau\Sigma)^{-1}$.

The third method for specifying the variance of the views is described in [Idzorek, 2004]. His method allows the specification of the view confidence in terms of the % move of the weights from no views to total certainty in the view. We will look at Idzorek's algorithm in the section on extensions.

## Applying Bayes Theorem

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 (4), (12) and (13) we can apply Bayes Theorem and derive our formula for the posterior distribution of asset returns.

Substituting (12) and (13) into (4) we have the following distribution

(15) $P(A|B) \sim N([(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q][(\tau\Sigma)^{-1} + P^T\Omega^{-1}P]^{-1},((\tau\Sigma)^{-1} + P^T\Omega^{-1}P)^{-1})$.

This is sometimes referred to as the Black-Litterman master formula. Two different derivations of this formula are in Appendix A. An alternate representation of the same formula for the mean returns (E(r)) and covariance (M) is

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

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

The derivation of formula (16) is shown in Appendix C. Remember that M, the posterior variance, is the variance of the posterior mean about the true, or population mean. It is essentially the uncertainty in the posterior mean, 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 (7). This is mentioned in [He and Litterman, 1999] but not in any of the other papers.

(18) $\Sigma_p = \Sigma + M$

Substituting the posterior variance from (17) we get

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

In the absence of views this reduces to

(19) $\Sigma_p = \Sigma + (\tau\Sigma) = (1 + \tau)\Sigma$

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]. [Idzorek, 2004] does not compute a new posterior variance, but instead just uses the known input variance of the returns about the mean. Further details on this topic are provided in the section on results.

There is some disagreement amongst practitioners about whether it is better to just use the known variance of returns, Σ, or whether to use an updated posterior variance of returns of the form show in formula (18). The updated posterior variance will be lower than either the prior or conditional variance, indicating that the addition of more information will reduce the uncertainty of the model. The variance from formula (18) 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. If M is fixed at zero as Idzorek does, then the posterior variance of returns equals the prior variance of returns. Given that there is some uncertainty in this value (M), then formula (18) provides a better estimator of the variance of returns than the prior variance of returns.

Since we are building the covariance matrix, Σ, from historical data we can compute τ from the number of samples. We can also estimate τ based on our confidence in the prior distribution. Note that both of these techniques provide some intuition for selecting a value of τ which is closer to 0 than to 1. [He and Litterman, 1999] and [Idzorek, 2004] both indicate that in their calculations they used small values of τ, on the order of 0.025 – 0.050. [Satchell and Scowcroft, 2000] state that many investors use a τ around 1 which does not seem to have any intuitive connection here. The next section of the document will provide some details on the impact of the value of τ.

[Mankert, 2006] derives the Black-Litterman 'master formula' using a Sampling Theory approach which yields τ as a ratio of the confidence in the prior to the confidence in the sampling distribution. She also provides a full derivation of formula (16) from formula (15).

Her argument about τ is interesting, and given how Black-Litterman is applied in the literature (setting $\Omega = \text{diag}(P(\tau\Sigma)P^t)$) it is reasonable. If we start from an equivalent point in Bayesian theory we see her argument 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 (20) by the same logic.

Given

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

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

then

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

Given this form, and relating back to formula (15) we can see that while $\Sigma = S_p$ and thus $\tau\Sigma = S_p/n$, her assumption that $\Omega = S_c$ is not correct. In fact, $\Omega = S_c/m$, and thus τ cannot be the ratio of n to m. This confusion arises from two issues. The first is the pragmatic approach chosen by most implementors of Black-Litterman where they relate Ω to τΣ because of the difficulty in independently estimating Ω in any reasonable way. The second and deeper issue is that while Σ is the variance of the prior returns, τΣ is the variance of the mean estimate about the true mean, or the square of the standard error of the estimate of Π. Ω is the square of the standard error of the estimates of the mean from the conditional distribution. There is no description within Black-Litterman, or anywhere within this framework for any parameter to represent the variance of the conditional returns about their mean. It seems this subtlety is missed by many authors.

Given formula (16) it is easy to let $\Omega \rightarrow 0$ showing that the return under 100% certainty of the views is

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

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

Finding an analytically tractable way to express and compute the posterior variance under 100% certainty is a challenging problem. Formula (4) 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 (6) is

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

If $\Omega \rightarrow 0$ (total confidence in views, and every asset is in at least one view) then formula (22) can be reduced to M = 0. If on the other hand the investor is not confident in their views, $\Omega \rightarrow \infty$, then formula (22) can be reduced to M = τΣ.

[Meucci, 2005] describes the transformation from (6) to (22), but does not show the full derivation. I have included that derivation in Appendix B.

### *The Impact of τ*

This section of the document will discuss the impact of the parameter τ.

We can divide the authors on two axis in terms of how they deal with τ. First, there is the camp that thinks τ should be near 1 which includes Satchell and Scowcroft, Meucci and others. Second, there is the camp that thinks τ should be a small number on the order of 0.025-0.05 including He and Litterman, Black and Litterman and Idzorek. For the most part, those authors in the first camp also take the posterior variance of returns to be the prior variance of returns. This seems somewhat consistent, if at odds with some of the theory behind the formulas. The authors in the second camp, with the exception of Idzorek use a new posterior variance, updated by the conditional distribution, and strictly larger than the prior variance of returns.

As an exercise to understand the impact of $\tau$ on the results, note that if we set $\Omega = P(\tau\Sigma)P^t$ as Meucci does and we assume P is of full rank, then we can substitute this into formula (15) as

$$
\begin{aligned}
E(r) &= \Pi + \tau\Sigma P^T [(P\tau\Sigma P^T) + \Omega]^{-1}[Q - P\Pi] \\
&= \Pi + \tau\Sigma P^T [(P\tau\Sigma P^T) + P(\tau\Sigma)P^t]^{-1}[Q - P\Pi] \\
&= \Pi + \tau\Sigma P^T [2(P\tau\Sigma P^T)]^{-1}[Q - P\Pi] \\
&= \Pi + (1/2)\tau\Sigma P^T(P^T)^{-1}(\tau\Sigma)^{-1}P^{-1}[Q - P\Pi] \\
&= \Pi + (1/2)P^{-1}[Q - P\Pi] \\
(23) \qquad &= \Pi + (1/2)[P^{-1}Q - \Pi]
\end{aligned}
$$

Our starting point was not standard with how Black-Litterman is usually implemented, but the result is interesting none the less. Clearly, deriving $\Omega$ from the variance of the prior distribution leads to eliminating $\tau$ from the final formula for E(r).

We can see a similar result if we substitute $\Omega = P(\tau\Sigma)P$ into formula (22).

$$
\begin{aligned}
M &= \tau\Sigma - \tau\Sigma P^T [P\tau\Sigma P^T + \Omega]^{-1}P\tau\Sigma \\
&= \tau\Sigma - \tau\Sigma P^T [P\tau\Sigma P^T + P(\tau\Sigma)P^t]^{-1}P\tau\Sigma \\
&= \tau\Sigma - \tau\Sigma P^T [2P\tau\Sigma P^T]^{-1}P\tau\Sigma \\
&= \tau\Sigma - (1/2)\tau\Sigma P^T(P^T)^{-1}(\tau\Sigma)^{-1}(P^T)^{-1}P\tau\Sigma \\
&= \tau\Sigma - (1/2)(\tau\Sigma) \\
(24) \qquad &= (1/2)(\tau\Sigma)
\end{aligned}
$$

In both cases, our choice for $\Omega$ 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, and the variance of the estimate will be cut in half.

If we instead solve for the more useful general case of $\Omega = \alpha P(\tau\Sigma)P^t$ where $\alpha \geq 1$, substituting into (16) and following the same logic as used to derive (24) we get

$$(25) \qquad = \Pi + (1/(1+\alpha))[P^{-1}Q - \Pi]$$

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 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.

## 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 is based on formula (10), but rather than solving for $\Pi$, solve for w.

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

As a first test of our algorithm we verify that when the investor has no views that the weights are correct, substituting formula (19) into (26) we get

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

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

Given this result, it is clear that the output weights with no views will be impacted by the choice of $\tau$, but only if the [He and Litterman, 1999] form of the posterior variance is used. He and Litterman are the only authors which mention this fact. They 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 (27).

### *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 also one of the most authentic sources on the Black-Litterman model, as Robert Litterman co-authored both this paper and the original paper [Black and Litterman, 1992].

[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 posterior variance as calculated in formulas (22) and (18) as the input variance for the optimization step.

Appendix D contains a complete description of the algorithm used including all the formulas and calculations required.

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

| Asset | $P_0$ | $P_1$ | μ | $w_{eq}/(1+\tau)$ | w* | w* - $w_{eq}/(1+\tau)$ |
|---|---|---|---|---|---|---|
| Australia | 0.0 | 0.0 | 4.3 | 16.4 | 1.5% | .0% |
| Canada | 0.0 | 1.0 | 8.9 | 2.1% | 53.9% | 51.8% |
| France | -0.295 | 0.0 | 9.3 | 5.0% | -.5% | -5.4% |
| Germany | 1.0 | 0.0 | 10.6 | 5.2% | 23.6% | 18.4% |
| Japan | 0.0 | 0.0 | 4.6 | 11.0% | 11.0% | .0% |
| UK | -0.705 | 0.0 | 6.9 | 11.8% | -1.1% | -13.0% |
| USA | 0.0 | -1.0 | 7.1 | 58.6% | 6.8% | -51.8% |
| q | 5.0 | 4.0 | | | | |
| ω/τ | .043 | .017 | | | | |
| λ | .193 | .544 | | | | |

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 a Black-Litterman model which leaves Σ, the known variance of the returns from the prior distribution, as the variance of the posterior return estimates This is a significant difference from the algorithm used in [He and Litterman, 1992], 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 'Un-modified posterior variance mode' of the algorithm. This is what I believe is the method used by Idzorek.

Table 2 – He and Litterman version of Black-Litterman Model with Idzorek data.

| Asset Class | $\mu$ | $w_{eq}$ | w* | He-Litterman Mode | Idzorek |
|---|---|---|---|---|---|
| US Bonds | .07 | 18.87% | 28.96% | 10.09% | 10.54 |
| Intl Bonds | .50 | 25.49% | 15.41% | -10.09% | -10.54 |
| US LG | 6.50 | 11.80% | 9.27% | -2.52% | -2.73 |
| US LV | 4.33 | 11.80% | 14.32% | 2.52% | -2.73 |
| US SG | 7.55 | 1.31% | 1.03% | -.28% | -0.30 |
| US SV | 3.94 | 1.31% | 1.59% | .28% | 0.30 |
| Intl Dev | 4.94 | 23.59% | 27.74% | 4.15% | 3.63 |
| Intl Emg | 6.84 | 3.40% | 3.40% | .0% | 0 |

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

Table 3 – Idzorek version of the Black-Litterman Model with Idzorek data

| Country | $\mu$ | $w_{eq}$ | w | Idzorek Mode | Idzorek |
|---|---|---|---|---|---|
| US Bonds | .07 | 19.34% | 29.89% | 10.55% | 10.54 |
| Intl Bonds | .50 | 26.13% | 15.58% | -10.55% | -10.54 |
| US LG | 6.50 | 12.09% | 9.37% | -2.72% | -2.73 |
| US LV | 4.33 | 12.09% | 14.81% | 2.72% | -2.73 |
| US SG | 7.55 | 1.34% | 1.04% | -.30% | -0.30 |
| US SV | 3.94 | 1.34% | 1.64% | .30% | 0.30 |
| Intl Dev | 4.94 | 24.18% | 27.77% | 3.59% | 3.63 |
| Intl Emg | 6.84 | 3.49% | 3.49% | .0% | 0 |

***Matching the Results of Fusai and Meucci***

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.

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

The Mahalanobis distance is distributed as chi-square with n degrees of freedom (n is the number of assets). Thus the probability of this event occuring can be computed as:

(29) $P(q) = 1 – F(M(q))$

Where F(M(q)) is the CDF of the chi square 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

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

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

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

The primary change in the Black-Litterman model that they introduce is that $\Omega = \alpha P\tau\Sigma P^T$ 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. They also fix the parameter $\tau$ at 1, which means that $\Sigma$ is directly the variance of the estimate. 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. Unfortunately, if we set $\tau$ to a value similar to what [He and Litterman, 1999] use in formulas (28) or (30) we quickly find that a small value of $\tau$ renders very extreme values of the Mahalanobis distance and not entirely useful values for the sensitivities.

The reason for the problems caused by $\tau \neq 1$ is that views which are well within the variance of the returns ($\Sigma$) and thus have a small Mahalanobis distance, when compared to ($\tau\Sigma$) the variance of the estimate are very excessive. This fits with our intuition, if the investors views did not significantly exceed the variance of the estimate they would not be interesting.

### ***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 very interesting to understand what value they use for a posterior variance when they set $\tau = 1$. They likely take the same approach as that taken by Idzorek and Meucci.

[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. As an alternative I formulated the problem in terms of what I call mini-portfolios which represent the various partially hedged assets either separately or by country, this problem requires additional work up front to calculate the returns and variances of the mini-portfolios, but does not require any constraints on the reverse optimization. Unfortunately, this approach does not work with their data to generate their results. It has been used with the [He and Litterman ,1999] data to validate the method. Most likely the problem is related to rounding of values in the covariance matrix, or perhaps even an error in the covariance matrix as printed in the article. 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] and in [Krishnan and Mains, 2006]. [Idzorek, 2004] presents a means to calibrate the confidence or variance of the investors views. [Krishnan and Mains, 2006] present a method to incorporate additional factors into the model.

### *Idzorek's Extension*

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. formula (30) shows this model.

(31) $\%\ \text{confidence} = (w^* - w_{mkt}) / (w_{mk} - w_{100})$

$w_{100}$ 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.

Thus, if the user specifies a confidence value for the view we can perform a line search to fit formula (31) for each view. Idzorek recommends the use of a least squares fit between the weights computed at each step and the target weights. As I implemented this extension, a simple line search can be used to find the proper variance of the view in order to match the confidence level specified. The result of this line search will be values of $\omega_i$ for the individual views. Once $\Omega$ has been computed using the method for each view, then all the views can be rolled up 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.

***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.

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 ($\Omega = 0$) 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 $\Omega$ which is $diag(P(\tau\Sigma)P^t)$. 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 $\Omega$ value for the desired confidence level.

Table 4A – Results for View 1

| Asset | $\omega_1$ | $w_{mkt}$ | $w^*$ | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| Intl Dev Equity | .000708875 | 24.18% | 27.77% | 29.28% | 70.47% |

Table 4B – Calibrated Results for View 1

| Asset | $\omega_1$ | $w_{mkt}$ | $w^*$ | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| Intl Dev Equity | .002124023 | 24.18% | 25.46% | 29.28% | 25.02% |

Table 5A – Results for View 2

| Asset | $\omega_2$ | $w_{mkt}$ | $w^*$ | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| US Bonds | .000140650 | 19.34% | 29.89% | 38.78% | 54.26% |
| Intl Bonds | .000140650 | 26.13% | 15.58% | 6.69% | 54.26% |

Table 5B – Calibrated Results for View 2

| Asset | $\omega_2$ | $w_{mkt}$ | w* | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| US Bonds | .000140667 | 19.34% | 29.06% | 38.78% | 50.00% |
| Intl Bonds | .000140667 | 26.13% | 16.41% | 6.69% | 50.00% |

Table 6A - Results for View 3

| Asset | $\omega_3$ | $w_{mkt}$ | w* | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| US LG | .000865628 | 12.09% | 9.37% | 8.09% | 68.06% |
| US LV | .000865628 | 12.09% | 14.81% | 16.09% | 68.06% |
| US SG | .000865628 | 1.34% | 1.04% | .90% | 68.06% |
| US SV | .000865628 | 1.34% | 1.64% | 1.78% | 68.06% |

Table 6B – Calibrated Results for View 3

| Asset | $\omega_3$ | $w_{mkt}$ | w* | $w_{100\%}$ | Implied Confidence |
|---|---|---|---|---|---|
| US LG | .000466003 | 12.09% | 9.49% | 8.09% | 65.01% |
| US LV | .000466003 | 12.09% | 14.69% | 16.09% | 65.01% |
| US SG | .000466003 | 1.34% | 1.05% | .90% | 65.01% |
| US SV | .000466003 | 1.34% | 1.63% | 1.78% | 65.01% |

Then we use the freshly computed values for the $\Omega$ matrix with all views specified together and arrive at the following final result blending all 3 views together.

Table 7 – Final Results for Idzorek's Confidence Extension Example

| Asset | View 1 | View 2 | View 3 | μ | σ | $w_{mkt}$ | Posterior Weight | change |
|---|---|---|---|---|---|---|---|---|
| US Bonds | 0.0 | -1.0 | 0.0 | .1 | 3.2 | 19.3% | 29.6% | 10.3% |
| Intl Bonds | 0.0 | 1.0 | 0.0 | .5 | 8.5 | 26.1% | 15.8% | 10.3% |
| US LG | 0.0 | 0.0 | 0.9 | 6.3 | 24.5 | 12.1% | 8.9% | 3.2% |
| US LV | 0.0 | 0.0 | -0.9 | 4.2 | 17.2 | 12.1% | 15.2% | 3.2% |
| US SG | 0.0 | 0.0 | 0.1 | 7.3 | 32.0 | 1.3% | 1.0% | -.4% |
| US SV | 0.0 | 0.0 | -0.1 | 3.8 | 17.9 | 1.3% | 1.7% | .4% |
| Intl Dev | 1.0 | 0.0 | 0.0 | 4.8 | 16.8 | 24.2% | 26.0% | 1.8% |
| Intl Emg | 0.0 | 0.0 | 0.0 | 6.6 | 28.3 | 3.5% | 3.5% | -.0% |
| Total | | | | | | | 101.8% | |
| Return | 5.2 | .2 | 2.0 | | | | | |
| Omega/ tau | .08496 | .00563 | .01864 | | | | | |
| Lambda | .002 | -.006 | -.002 | | | | | |

***Two-Factor Black-Litterman***

[Krishnan and Mains, 2005] developed an extension to the Black-Litterman 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 (9), but add an additional term for the new market factor(s).

(32) $$U = w^T \Pi - \left(\frac{\delta_0}{2}\right) w^T \Sigma w - \sum_{j=1}^{n} \delta_j w^T \beta_j$$

U is the investors utility, this is the objective function during portfolio optimization.

w is the vector of weights invested in each asset
Π is the vector of equilibrium excess returns for each asset
Σ 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 (32) we can take the first derivative with respect to w in order to solve for the equilibrium asset returns.

(33) $$\Pi = \delta_0 \Sigma w + \sum_{j=1}^{n} \delta_j \beta_j$$

Comparing this to formula (10), the simple reverse optimization formula, we see that the equilibrium excess return vector (Π) is a linear composition of (10) 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 δ 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, $v_j$, such that $v_j^T \beta_j = 0$. 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 δ by multiplying formula (33) 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 ≥ 1 we can multiply formula (33)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 $\delta_j$.

$$\delta_j = \frac{(r_j - \delta_0 v_j^T \Sigma v_j)}{(v_j^T \beta_j)}$$

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 (32) we are back to the standard utility function

$$U = w^T \hat{\Pi} - (\frac{\delta_0}{2}) w^T \Sigma w$$

and from formula (13)

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

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

$$\textit{thus}$$

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

We can directly substitute $\hat{\Pi}$ and $\hat{Q}$ into formula (16) 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 and set $\tau = 1$. They actually ignore $\tau$ in their paper which is the same as setting it to 1. They do not adjust the posterior variance, but set it to the known prior variance of the returns as most other authors do. They provide all of the data needed to reproduce their results given the set of formulas in this section.

## *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. 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 papar, [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 $[P^T\Omega^{-1}P]((\tau\Sigma)^{-1} + P^T\Omega^{-1}P)^{-1}$. This represents the weight of the view returns in the mixing. As $\Omega \rightarrow 0$, then the weight on views $\rightarrow$ 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 Jagnannathan, 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 Jagnannathan, 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 $\tau$, 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 with links to many of these papers. I used [DeGroot, 1970] as my reference on Bayesian Statistics, there are many references with the results mentioned in this paper.

[Bevan and Winkelmann, 1998] Using the Black-Litterman Global Asset Allocation Model: Three Years of Practical Experience, Bevan and Winkelmann, June 1998. Goldman Sachs Fixed Income Research paper.

[Black and Litterman, 1991] Global Portfolio Optimization, Fischer Black and Robert Litterman, 1991. Journal of Fixed Income, 1991.

[Black and Litterman, 1992] Global Portfolio Optimization, Fischer Black and Robert Litterman, 1992. Financial Analysts Journal, Sept/Oct 1992.

[Christadoulakis, 2002] Bayesian Optimal Portfolio Selection: The Black-Litterman Approach, Christadoulakis, 2002. Class Notes.

[Da and Jagnannathan, 2005] Teaching Note on Black-Litterman Model, Da and Jagnannathan, 2005. Teaching notes.

[DeGroot, 1970] Optimal Statistical Decisions, 1970. Wiley Interscience.

[Frost and Savarino, 1986] An Empirical Bayes Approach to Efficient Portfolio Selection, Peter Frost and James Savarino, Journal of Financial and Quantitative Analysuis, Vol 21, No 3, September 1986.

[Fusai and Meucci, 2003] Assessing Views, Risk Magazine, **16**, 3, S18-S21.

[He and Litterman, 1999] The Intuition Behind Black-Litterman Model Portfolios, He and Robert Litterman, 1999. Goldman Sachs Asset Management Working paper.

[Herold, 2003] Portfolio Construction with Qualitative Forecasts, Ulf Herold, 2003. J. of Portfolio Management, Fall 2003, p61-72.

[Idzorek, 2004] A Step-By-Step guide to the Black-Litterman Model, Incorporating User-Specified Confidence Levels, Idzorek, 2004. Working paper.

[Koch, 2005] Consistent Asset Return Estimates. The Black-Litterman Approach. Cominvest presentation.

[Krishnan and Mains, 2005], The Two-Factor Black-Litterman Model, Hari Krishnan and Norman Mains, 2005. July 2005, Risk Magazine.

[Litterman, et al, 2003] Beyond Equilibrium, the Black-Litterman Approach, Litterman, 2003. Modern Investment Management: An Equilibrium Approach by Bob Litterman and the Quantitative Research Group, Goldman Sachs Asset Management, Chapter 7.

[Mankert, 2006] The Black-Litterman Model – Mathematical and Behavioral Finance Approaches Towards its Use in Practice, Mankert, 2006. Licentiate Thesis.

[Meucci, 2005] Risk and Asset Allocation, 2005. Springer Finance.

[Meucci, 2006] Beyond Black-Litterman in Practice: A Five-Step Recipe to Input Views on non-Normal Markets, Attilio Meucci, Working paper.

[Qian and Gorman, 2001], Conditional Distribution in Portfolio Theory, Edward Qian and Stephen Gorman, Financial Analysts Journal, September, 2001.

[Satchell and Scowcroft, 2000] A Demystification of the Black-Litterman Model: Managing Quantitative and Traditional Portfolio Construction, Satchell and Scowcroft, 2000, J. of Asset Management, Vol 1, 2, 138-150..

[Theil and Goldberger, 1963].

## *Appendix A*

This appendix includes the two most common derivations of the Black-Litterman master formula. The first is based on Theil's Mixed Estimation approach which is based on Generalized Least Squares. The second is 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.

### *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 μ is the mean of the conditional distribution and v is the normally distributed residual with mean 0 and variance Σ.

Both S and Σ 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 $$\hat{E}(r) = \left[ \begin{bmatrix} I & I \end{bmatrix} \begin{bmatrix} S/n & 0 \\ 0 & \Sigma \end{bmatrix}^{-1} \begin{bmatrix} I' \\ I' \end{bmatrix} \right]^{-1} \begin{bmatrix} I' & I' \end{bmatrix} \begin{bmatrix} S/n & 0 \\ 0 & \Sigma \end{bmatrix}^{-1} \begin{bmatrix} x \\ \mu \end{bmatrix}$$

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

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

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

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

A.7 $$\hat{E}(r)-E(r)=\left[(S/n)^{-1}+\Sigma^{-1}\right]^{-1}\left[u(S/n)^{-1}+v\Sigma^{-1}\right]$$

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

$$E(\hat{E}(r)-E(r))^2=(\left[(S/n)^{-1}+\Sigma^{-1}\right]^{-1}\left[u(S/n)^{-1}+v\Sigma^{-1}\right])^2$$

$$E(\hat{E}(r)-E(r))^2=\left[(S/n)^{-1}+\Sigma^{-1}\right]^{-2}\left[u^2(S/n)^{-2}+v^2\Sigma^{-2}+uv(S/n)^{-1}\Sigma^{-1}\right]$$

We know from above that $E(uu')=(S/n)$ , $E(vv')=\Sigma$ and $E(uv')=0$ because u and v are independent variables, so substituting

$$E(\hat{E}(r)-E(r))^2=\left[(S/n)^{-1}+\Sigma^{-1}\right]^{-2}\left[(S/n)(S/n)^{-2}+\Sigma\Sigma^{-2}+0\right]$$

A.8 $$E(\hat{E}(r)-E(r))^2=\left[(S/n)^{-1}+\Sigma^{-1}\right]^{-1}$$

***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

A.9 $$\xi(x)\propto \exp\left((S/n)^{-1}(E(r)-x)^2\right)$$

Next, we consider the PDF for the conditional distribution.

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

So ξ(μ|x) the PDF of P(B|A) satisfies

A.10 $\xi(\mu|x) \propto \exp\left(\Sigma^{-1}(E(r)-\mu)^2\right)$

Substituting A.9 and A.10 into formula (1) from the text, we have an expression which the PDF of the posterior distribution will satisfy.

A.11 $\xi(x|\mu) \propto \exp\left(-\left(\Sigma^{-1}(E(r)-\mu)^2+(S/n)^{-1}(E(r)-x)^2\right)\right)$ ,

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

Considering only the quantity in the exponent and simplifying

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

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

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

If we introduce a new term y, where

A.12 $y=\dfrac{\left(\mu\Sigma^{-1}+x(S/n)^{-1}\right)}{\left(\Sigma^{-1}+(S/n)^{-1}\right)}$

and then substitute in the second term

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

Then add $0=y^2(\Sigma^{-1}+(S/n)^{-1})-(\mu\Sigma^{-1}+x(S/n)^{-1})^2(\Sigma^{-1}+(S/n)^{-1})^{-1}$

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

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

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

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

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

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

 

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

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

A.13 $$\xi(x|\mu) \propto \exp\left(-\left[\left(\Sigma^{-1} + (S/n)^{-1}\right)^{-1}(E(R) - y)^2\right]\right)$$

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

A.14 $$\left(\Sigma^{-1} + (S/n)^{-1}\right)^{-1}$$

## *Appendix B*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= (\Omega^{-1}P)[(P^{-1}\Omega)((P\tau\Sigma)^{-}(\tau\Sigma)^{-1}P^{-1}\Omega + (P\tau\Sigma)^{-}P^T)^{-11}(P\tau\Sigma)]$$

$$= (\Omega^{-1}P)[(P^{-1}\Omega)(\Omega + P\tau\Sigma P^T)^{-11}(P\tau\Sigma)]$$

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

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

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

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

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

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

## *Appendix C*

This appendix presents a derivation of the alternate formulation of the Black-Litterman master formula for the posterior expected return. Starting from formula (15) we will derive formula (16).

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

Separate the parts of the second term

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

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

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

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

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

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

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

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

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

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

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

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

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

## *Appendix D*

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,

Σ Matrix of covariances between the asset classes. Can be computed from historical data.

$r_f$ Risk free rate for base currency

δ 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.

τ 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, Π using formula (10).

(10) $$\Pi = \delta \Sigma w$$

Then we formulate the investors views, and specify P, Ω, and Q. Given k views and n assets, then P is a k × n matrix where each row sums to 0 (relative view) or 1 (absolute view). Q is a k × 1 vector of the excess returns for each view. Ω is a diagonal k × 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 $p^T \tau \Sigma_i p$ (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 (16).

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

This following two steps are not needed to match the algorithm used by Idzorek.

Then we must compute the posterior variance using formula (22).

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

Closely followed by the computation of the sample variance from formula (18).

(18) $$\Sigma_p = \Sigma + M$$

And now we can compute the portfolio weights for the optimal portfolio on the unconstrained efficient frontier from formula (26).

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

 