The Black-Litterman Model In Detail

Initial: February 2007
Revised: June 20, 2014

jwalters@blacklitterman.org

Abstract[1]

In this paper we survey the literature on the Black-Litterman model. This survey is provided both as a chronology and a taxonomy as there are many claims on the model in the literature. We provide a complete description of the canonical model including full derivations from the underlying principles using both Theil's Mixed Estimation model and Bayes Theory. The various parameters of the model are considered, along with information on their computation or calibration. Further consideration is given to several of the key papers, with worked examples illustrating the concepts.

1 The author gratefully acknowledges feedback and comments from Attilio Meucci and Boris Gnedenko.

Introduction

The Black-Litterman Model was first published by Fischer Black and Robert Litterman in 1990. During the twenty plus years since the original papers, many authors have published research referring to their model as Black-Litterman. This has lead to a variety of models being labeled as Black-Litterman even though they may be very different from the original model created by Black and Litterman. In the chronological survey of the literature we will introduce several papers which make significant contributions to the literature on the Black-Litterman model. We will also document the taxonomy of models which have been labeled as Black Litterman. We will further refer to the Black-Litterman model as described in the papers by the original authors as the canonical Black-Litterman model.

The canonical Black-Litterman model makes two significant contributions to the problem of asset allocation. First, it provides an intuitive prior, the 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 equilibrium market portfolio as a starting point for shrinkage is a significant improvement to the process of return estimation.

Second, the Black-Litterman model provides a clear way to specify investors views on returns 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 outputs estimates of expected excess returns and the estimates corresponding precision Prior to Black and Litterman, (1991a), nothing similar had been published for the portfolio selection problem. Theil's mixing model had been developed, but nobody had applied it to the problem of estimating asset 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.

The state of the art in the portfolio selection problem has changed significantly during the time since the publication of the Black-Litterman model. Because of its rich theoretical basis it can be applied alongside many modern portfolio selection approaches as can be seen in the literature..

The rest of the document is structured in the following manner. First we will address the literature survey and taxonomy, then we will describe each of the reference models, and then the estimation model. Finally we will illustrate the concepts from many of the papers described in the literature survey.

Historical Taxonomy and Literature Survey

This section of the paper will provide survey of the literature and will classify the model used by each of the various authors. The primary dimensions we will use to classify the authors' models will be; does it specify the estimates as distributions or as point estimates, and secondly does it include the parameter τ. In order for an author's model to be canonical Black-Litterman it would need to match both these conditions.

We will collect the models into three distinct Reference Models based on our dimensions provided above. We describe the situation where both conditions are met as the Canonical Reference Model. This is the model described by the original authors. The Alternative Reference Model describes models which use point estimates, but for some reason include τ which now becomes a scaling factor with no quantitative basis. These authors treat the model as just a shrinkage/mixing model and lose the connection to Bayesian statistics. As we will see, using the Bayes law formula, but substituting in different variables is not theoretically tenable.. Finally the Beyond Black-Litterman Reference Model uses point estimates and no longer includes τ at all. Here we see a pure shrinkage/mixing model. Each of the Reference Models is described in depth later in the paper.

The Black-Litterman model was first published by Fischer Black and Robert Litterman of Goldman Sachs in an internal Goldman Sachs Fixed Income Research Note, Black and Litterman (1990). This research note was extended into a paper and published in the Journal of Fixed Income in 1991, Black and Litterman (1991a). This paper does not provide all the formulas used in the model, but does provide a good overview of the features of the model. A second internal Goldman Sachs Fixed Income Research Note was published the same year, Black and Litterman (1991b). This second paper was later extended and published in the Financial Analysts Journal (FAJ) as Black and Litterman (1992). 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 based on the global equilibrium, see Litterman (2003) for more details on the methods required to solve this problem. Unfortunately, because of the merging of the two problems, their results are difficult to reproduce. He and Litterman (1999) is the last paper by one of the original authors and it does provide more detail on the workings of the model, but not quite a complete set of formulas. They do provide a much simpler to reproduce working example.

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) attempted to demystify the Black-Litterman model, but instead introduced a new non-Bayesian expression of the model. Their model uses point estimates for the prior and the views, and uses τ and Ω only to control the amount of shrinkage of the views onto the prior. Because they use point estimates instead of distributions, their model does not include any information on the precision of the estimate. This allows them to recommend setting $\tau = 1$. They also introduce a stochastic τ, but because they use point estimates this really becomes a model with a stochastic covariance of returns. Th model was used occasionally in the literature after this point, but was largely replaced by Meucci's model in the mid-2000's.

Drobetz (2001) provides further description of the Black-Litterman model including a good discussion of how to interpret the confidence in the estimates including a diagram. He works an example. [Need more text here]

Fusai and Meucci (2003) introduced yet another non-Bayesian variant of the model which removed the parameter τ all together. Meucci (2005) followed up on this paper and coined the phrase, "Beyond Black-Litterman". Once the model is viewed as only a shrinkage model, Ω alone provides enough degrees of freedom to the shrinkage and τ is superfluous and confusing. Since the mid-2000's we have seen a mixture of the canonical and "Beyond Black-Litterman" model used in the literature[2]. Since Fuasi and Meucci (2003) there has been little use of the hybrid reference model until Michaud, et al (2013) debunked it.

Firoozye and Blamont (2003), provide a good overview of the Canonical Reference Model and attempt to provide intuition in the setting of the parameter, τ. They illustrate the reduction in variance of the posterior estimate as a result of the mixing.

Herold (2003) provides an alternative view of the problem where he examines optimizing alpha generation for active portfolio management, essentially specifying that the sample distribution has zero mean. He uses the alternative reference model with point estimates and tracking error to determine how much shrinkage to allow. The two significant contributions of his paper are; the application of the model to active portfolio management and some additional measures which can be used to validate that the views are reasonable.

Koch (2004) 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). He does include a slide on sensitivity of the posterior estimate on τ, but he uses the alternative reference model so this information is not valid for the canonical model.

Idzorek (2005) introduced a technique for specifying Ω such that the impact of the shrinkage was specified in terms of percentage of change in the weights between 0 and the change caused by 100% confidence. His paper uses the Alternative Reference Model, but his technique can be applied to the Canonical Black-Litterman model because it is sensitive to the value of τ specified by the investor.

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 detailed 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 approaches the problem from the point of view of sampling theory, and provides some new intuition about the value τ in the context of the alternative reference model.

Meucci (2006) provides a method to use non-normal views in Black-Litterman. Meucci (2008) extends this method to any model parameter, and allow for both analysis of the full distribution as well as scenario analysis.

Beach and Orlov (2006) illustrate using GARCH models to generate the views. They use a model of international equity across 20 developed countries. They show how the results change as the investor uses different values for τ. They do not provide exact details on the uncertainty of the views and appear to be using an alternative reference model even though their techniques can be applied to the

2 A more comprehensive literature survey can be found at http://www.blacklitterman.org/methods.html.

canonical reference model.

Braga and Natale (2007) describes a method of calibrating the uncertainty in the views using Tracking Error Volatility (TEV). They provide sensitivities for the posterior estimates to the various views. [Verify] This metric is a well known for it's use in active portfolio management. They use the Alternative Reference Model but their work could be applied to the Canonical Reference Model a well.

Martellini and Ziemann (2007) describe an approach to active management of a fund of hedge funds. They use VaR as their objective function for the reverse optimization and they use a variant of the CAPM model extended to include skewness and kurtosis in determining their neutral portfolio. They use a factor model to generate ankings and then convert the rankings into their confidence in the views. They do not use the Bayesian features of the model, but rather use point estimates and thus do not have information on the precision of their estimates.

Giacommeti, et al (2007) provide an approach to computing the neutral portfolio using stable paretian distributions rather than the normal distribution described in Black and Litterman (1992). They also use multiple different risk measures for the portfolio selection model; variance VaR, and CVaR,

Cheung (2009) intoduces the concept of an augmented model. This version of the model integrates a factor model and does a joint estimation of the factor returns. Cheung uses the Alternative Reference Model, but his variant of the model could easily be used with the Canonical Reference Model.

Bertsimas, et al (2013) They introduce a way to measure the alignment of the views with the prior estimate by comparing the view portfolio weights to the eigenvalues from the prior covariance matrix.

Michaud, et al (2013) provides a blistering critique of the Alternative Reference Model. It is indeed a mixture of fact and opinion, only statistical methods will yield valid prior estimates. Because of it's focus on the Alternative Reference Model (a significant part of the paper is devoted to problems with point estimates) much of it is not relevant to the Canonical Reference Model. Further, the basic arguments restrict the reader to only basic statistical properties of time series (essentially ignoring non-stationarity and auto-regressive properties) and ignores richer, state of the art econometric models.

There are some additional teaching notes put together through the years, Christodoulakis (2002) and Da and Jagnannathan (2005) which provide very basic ideas for class room discussions. The first describes the Canonical Reference model but does not explain what the posterior variance represents. The second quickly describes the model and then works an example illustrating some ways to arrive at the various inputs to the model. The authors use the Alternative Reference Model as they include tau, but do not compute a updated posterior covariance..

Much of the work presented with any of the reference models can be used with the Canonical Reference Model, even if it has been initially formulated for the Alternative Reference Model. Later in this paper we will show how many of these results can be generalized to work with the Canonical Reference Model.

The Canonical Black-Litterman Reference Model

The reference model for returns is the base upon which the rest of the Black-Litterman model 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, some authors of papers on the

Black-Litterman model use reference models, e.g. Alternative Reference Model or Beyond Black-Litterman, which do not have the same theoretical basis as the canonical one which was initially specified in Black and Litterman (1992).

We start with normally distributed expected returns

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

The fundamental goal of the Black-Litterman model is to model these expected returns, which are assumed to be normally distributed with mean μ and variance Σ. Note that we will need at least these values, the expected returns and covariance matrix later as inputs into a portfolio selection model.

We define μ, the unknown mean return, as a random variable itself distributed as

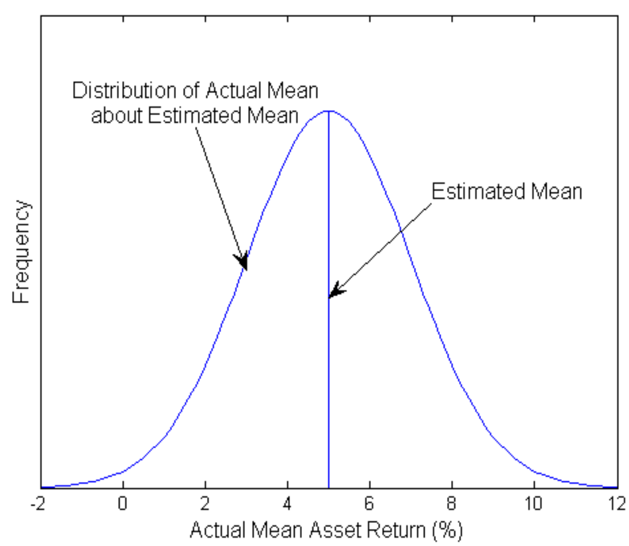
$\mu \sim N(\pi, \Sigma_\pi)$

π is our estimate of the mean and Σ_π is the variance of the unknown mean, μ, about our estimate. Another way to view this simple linear relationship is shown in the formula below.

(2) $\mu = \pi + \epsilon$

The prior returns are normally distributed around π with a disturbance value ε. The diagram below shows the distribution of the actual mean about the estimated mean of 5% with a standard deviation of 2% (or a variance of 0.0004). We can see that this is not a very precise estimate from the width of the peak.

Figure 1 - Distribution of Actual Mean about Estimated Mean



ε is normally distributed with mean 0 and variance Σ_π and is assumed to be uncorrelated with μ. We can complete the reference model by defining Σ_r as the variance of the returns about our estimate π. From formula (2) and the assumption above that ε and μ are not correlated, then the formula to compute Σ_r is

(3) $\Sigma_r = \Sigma + \Sigma_\pi$

Formula (3) 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. $\varepsilon \equiv 0$, then $\Sigma_r = \Sigma$. As our estimate gets worse, e.g. Σ_π increases, then Σ_r increases as well.

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

(4) $r \sim N(\pi, \Sigma_r)$

A common misconception about the canonical Black-Litterman reference model is that formula (1) is the reference model, and that μ is a point estimate. We have previously labeled this approach, the Alternate Reference Model. Several authors approach the problem from this point of view so we cannot neglect it, but it is a fundamentally different model. When considering results from various Black-Litterman implementations it is important to understand which reference model is being used in order to understand how the various parameters will impact the results.

Computing the Equilibrium Returns

The Black-Litterman model starts with a neutral equilibrium portfolio for the prior estimate of returns. The model relies on General Equilibrium theory to state that if the aggregate portfolio is at equilibrium, each sub-portfolio must also be at equilibrium. It can be used with any utility function which makes it very flexible. In practice most practitioners use the Quadratic Utility function and assume a risk free asset, and thus the equilibrium model simplifies to the Capital Asset Pricing Model (CAPM). The neutral portfolio in this situation is the CAPM Market portfolio. Some of the references listed have used other utility functions, most notably Bertsimas (2013), but others have consider CVaR and other measures of portfolio risk without applying the same theoretical basis. In order to preserve the symmetry of the model, the practioner should use the same utility function both to identify the neutral portfolio as well as in the portfolio selection area..

Here we will use the Quadratic Utility function, CAPM and unconsrained mean-variance because it is a well understood model.

Given our previous assumptions, the prior distribution for the Black-Litterman model is the estimated mean excess return from the CAPM market portfolio. The process of computing the CAPM equilibrium excess returns is straight forward.

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

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

Where

r_f The risk free rate.

r_m The excess return of the market portfolio.

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

α The residual, or asset specific (idiosyncratic) excess return.

Under CAPM the idiosyncratic risk associated with an asset is uncorrelated with that from other assets, and this risk can be reduced through diversification. Thus the investor is rewarded for taking systematic risk measured by β, but is not rewarded for taking idiosyncratic risk associated with α.

In the CAPM world, all investors should hold the same risky portfolio, the CAPM market portfolio. Because all investors hold risky assets only in the market portfolio, at equilibrium the market capitalization of the various assets will determine their weights in the market portfolio. The CAPM market portfolio has the maximum Sharpe Ratio[3] of any portfolio on the efficient frontier.

The investor can also invest in a risk free asset. This risk free asset has essentially a fixed positive return for the time period over which the investor is concerned. It is generally some tenor from the sovereign bond yield curve for the investors local currency. Depending on how the asset allocation decision will be framed this risk free asset can range from a 4 week Treasury Bill (1 month horizon) to a 20 year inflation protected bond (20 year horizon).

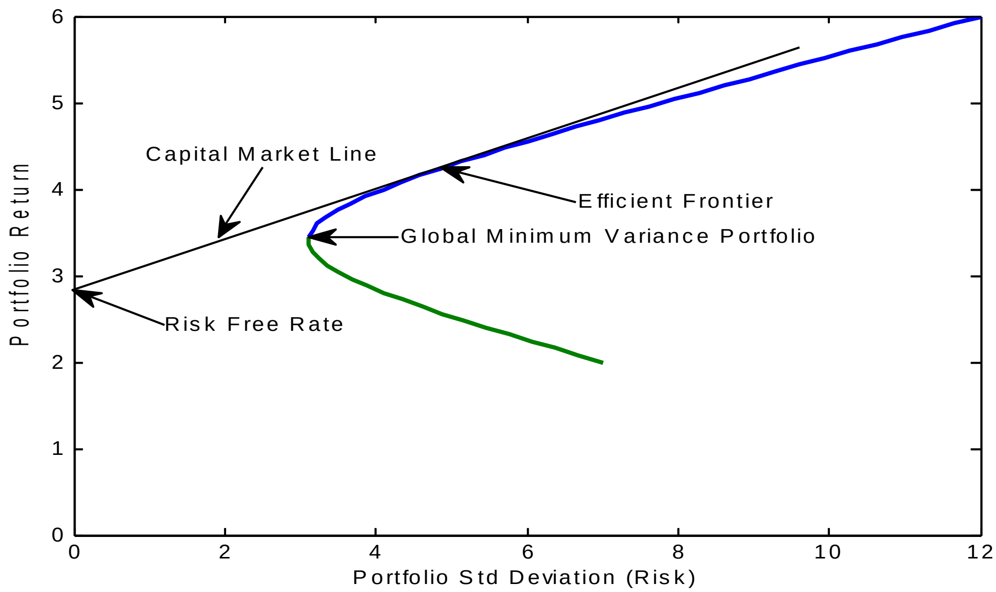
Note that the CAPM market portfolio contains all investable assets which makes it very hard to actually specify. Because we are in equilibrium, all sub-markets must also be in equilibrium so any sub-market we chose to use is part of the global equilibrium. While this allows us to reverse optimize the excess returns from the market capitalization and the covariance matrix, forward optimization from this point to identify the investors optimal portfolio within CAPM is problematic as we do not have information for the entire market portfolio. In general this is not actually the question an investor is asking, they usually select an investable universe and desire the optimal asset allocation within the universe, so the theoretical problem with the market portfolio can be ignored.

The Capital Market Line is a line through the risk free rate and the CAPM market portfolio. The Two Fund Separation Theorem, closely related to the CAPM, states that all investors should hold portfolios on the Capital Market Line. Any portfolio on the Capital Market Line dominates all portfolios on the Efficient Frontier, the CAPM market portfolio being the only point on the Efficient Frontier and on the Capital Market Line. Depending on their risk aversion an investor will hold arbitrary fractions of their wealth in the risk free asset and/or the CAPM market portfolio.

The diagram below illustrates the relationship between the Efficient Frontier and the Capital Market Line.

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

Figure 2 - Efficient Frontier



Since we are starting with the market portfolio, we will be starting with a set of weights which are all greater than zero and naturally sum to one. The market portfolio only includes risky assets, because by definition investors are rewarded only for taking on systematic risk. Thus, in the CAPM, the risk free asset with $\beta = 0$ will not be in the market portfolio. Later on we will see that our Bayesian investor may invest in the risk free asset based on their confidence in their return estimates.

We will constrain the problem by asserting that the covariance matrix of the returns, Σ, is known. In practice, this covariance matrix is estimated from historical return data. It is often computed from higher frequency data and then scaled up to the time frame required for the asset allocation problem. By computing it from actual historical data, we can ensure the covariance matrix is positive definite. Without basing the estimation process on actual data there are significant issues involved in ensuring the covariance matrix is positive definite. We could however apply shrinkage or random matrix theory filters to the covariance matrix in an effort to make it more robust.

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. Be aware 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

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

U	Investors utility, this is the objective function during Mean-Variance Optimization.
w	Vector of weights invested in each asset
Π	Vector of equilibrium excess returns for each asset

δ	Risk aversion parameter
Σ	Covariance matrix of the excess returns for the assets

U is a convex 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 (6) 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:

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

In order to use formula (7) to solve for the CAPM market portfolio, we need to have a value for δ, the risk aversion coefficient of the market. One way to find δ is by multiplying both sides of (7) by w^T and replacing vector terms with scalar terms.

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

Here our expression at equilibrium is that the excess return to the portfolio is equal to the risk aversion parameter multiplied by the variance of the portfolio.

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

r	Total return on the market portfolio ($r = w^T\Pi + r_f$)
r_f	Risk free rate
σ^{-2}	Variance of the market portfolio ($\sigma^2 = w^T\Sigma w$)

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 the example in their paper.

Given the Sharpe Ratio we can rewrite formula (8) for δ in terms of the Sharpe Ratio as

(9) $$\delta = \frac{SR}{\sigma_m}$$

Thus we can calibrate our returns in terms of formulas (8) or (9). As part of our analysis we must arrive at the terms on the right hand side of which ever formula we choose to use. For formula (8); this is r, r_f, and $\sigma^{2.}$ in order to calculate a value for δ. For formula (9) this is the Sharpe Ratio and σ.

In order to use formula (8) we need to have an implied return for the market portfolio which may be harder to estimate than the Sharpe Ratio of the market portfolio.

Once we have a value for δ, then we plug w, δ and Σ into formula (7) and generate the set of equilibrium asset returns. Formula (7) 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 (7) to yield the formula for the closed form calculation of the optimal portfolio weights in the absence of constraints.

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

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 ($\Sigma w = 1$) constraint. He illustrates how errors can be introduced during a reverse optimization process if constraints are assumed to be non-binding when they are actually binding for a given portfolio. Note that because we are dealing with the market portfolio which has only positive weights which sum to 1, we can safely assume that there are no binding constraints on the reverse optimization.

The only missing piece is the variance of our estimate of the mean. Looking back at the reference model, we need Σ_π. 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, $\Sigma_\pi = \tau\Sigma$, then the prior distribution P(A) is:

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

This is the prior distribution for the Black-Litterman model. It represents our estimate of the mean, which is expressed as a distribution of the actual unknown mean about our estimate.

Using formula (4) we can rewrite formula (11) in terms of Π as

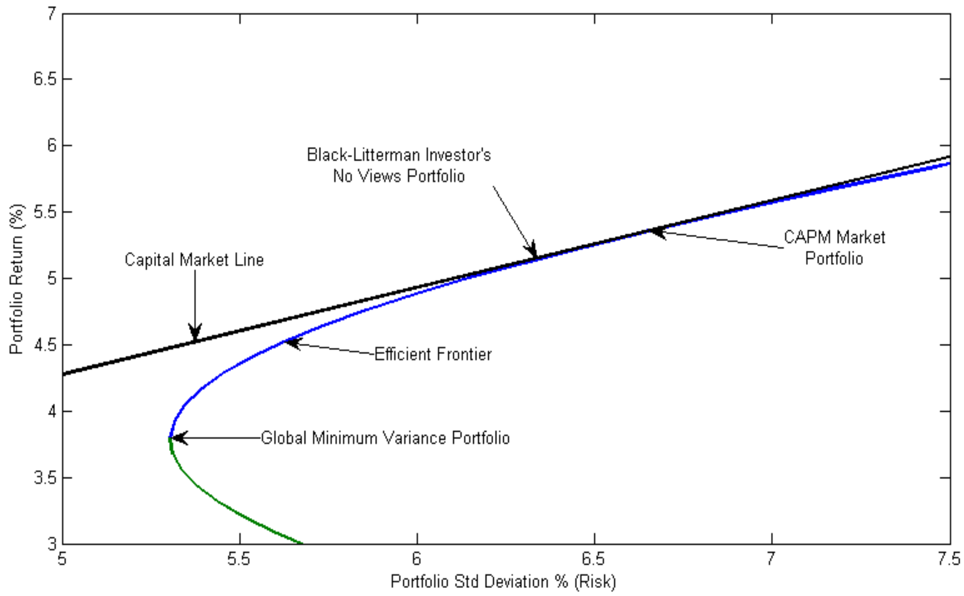
(12) $r_A \sim N(\Pi, (1+\tau)\Sigma)$

It is often written that an investor with no views and using an unconstrained mean-variance portfolio selection model will invest 100% in the neutral portfolio, but this is only true if they apply a budget constraint. Because of their uncertainty in the estimates they will invest $1/(1+\tau)$ in the neutral portfolio and $\tau/(1+\tau)$ in the risk free asset. We can see this is the case as follows, starting from (7):

$$\Pi = \delta\Sigma w$$
$$w = (\delta\Sigma)^{-1}\Pi$$
$$\hat{w} = ((1+\tau)\delta\Sigma)^{-1}\Pi$$
$$\hat{w} = (1/1+\tau)(\delta\Sigma)^{-1}\Pi$$
$$\hat{w} = (1/1+\tau)w$$

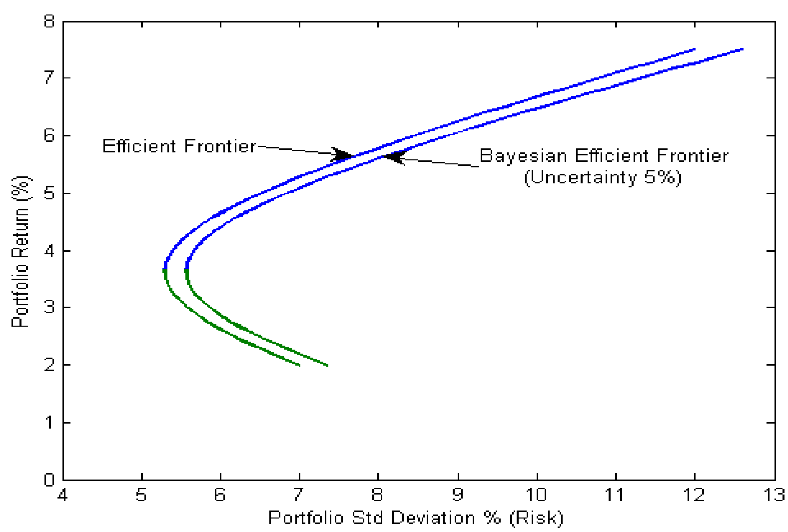
The diagram below illustrates this concept graphically.

Figure 3 - Investor's Portfolio in the Absence of Views



We can alternatively view the Bayesian efficient frontier as a shift to the right if we plot the efficient frontier generated with the increased covariance matrix and a budget constraint. In this case the uncertainty just pushes each point further to the right in risk/return space.

Figure 4 - Risk Adjusted Bayesian Efficient Frontier



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 any or all assets. In addition it is actually possible for the views to conflict, the mixing process will merge the views based on the confidence in the views and the confidence in the prior.

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

- P, a k×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, He and Litterman (1999) and Idzorek (2005) use a market capitalization weighed scheme, whereas Satchell and Scowcroft (2000) use an equal weighted scheme in their examples. In practice weights will be a mixture depending on the process used to estimate the view returns.
- Q, a k×1 vector of the returns for each view.
- Ω a k×k matrix of the covariance of the views. Ω is diagonal as the views are required to be independent and uncorrelated. Ω^{-1} is known as the confidence in the investor's views. The i-th diagonal element of Ω is represented as ω_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.

Ω 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 Ω may or may not be invertible. At a practical level we can require that $\omega > 0$ so that Ω is invertible, but we will reformulate the problem so that Ω is not required to be inverted.

As an example of how these matrices would be populated we will examine some investors views. Our example will have four assets and two views. First, a relative view in which the investor believes that Asset 1 will outperform Asset 3 by 2% with confidence. ω_1. Second, an absolute view in which the investor believes that Asset 2 will return 3% with confidence ω_2. 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:

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

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

$P(B|A) \sim N(Q, \Omega)$

We generally cannot convert this into a useful expression into asset space because of the mixture of relative and absolute views, and because the P matrix is not required to be of full rank. If we did express the views in asset space the formula is shown below.

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

This representation is not of any practical use. Incomplete views and relative views make the variance non-invertible, and relative views also impact the mean term. 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). It is however, interesting to see how the views are projected into the asset space.

Specifying Ω

Ω, 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 Ω.

There are several ways to calculate Ω.

- 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 along the weight dimension

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 follows:

(14)
$$\begin{aligned} \omega_{ij} &= p(\tau\Sigma)p^T \quad \forall\, i=j \\ \omega_{ij} &= 0 \quad \forall\, i \neq j \end{aligned}$$

or

$$\Omega = diag(P(\tau\Sigma)P^T)$$

This specification of the variance, or uncertainty, of the views essentially equally weights the investor's views and the market equilibrium weights. By including τ in the expression, the posterior estimate of the returns becomes independent of τ 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

$$\Omega = \frac{1}{c} P \Sigma P^t$$

He sets c >1, and one obvious choice for c is τ^{-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 specify the variance using a confidence interval around the estimated mean return, e.g. Asset has an estimated 3% mean return with the expectation it is 68% likely to be within the interval (2.0%,4.0%). Knowing that 68% of the normal distribution falls within 1 standard deviation of the mean allows us to translate this into a variance for the view of $(1\%)^2$.

Note that Ω is our uncertainty in the estimate of the mean, we are not specifying the variance of returns about the mean. This formulation of the variance of the view is consistent with the canonical reference model.

Use the Variance of Residuals from a Factor Model

If the investor is using a factor model to compute the views, they can 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:

(15) $$r=\sum_{i=1}^{n}\beta_i f_i+\epsilon$$

Where

r	is the return of the asset
β_i	is the factor loading for factor (i)
f_i	is the return due to factor (i)
ϵ	is an independent normally distributed residual

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

(16) $$V(r)=B\,V(F)B^T+V(\epsilon)$$

B	is the factor loading matrix
F	is the vector of returns due to the various factors

Given formula (15), and the assumption that ε is independent and normally distributed, then we can compute the variance of ε 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 (2005) 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. We choose to start with Theil's model because it is simpler and cleaner. We also work through the Bayesian version of the derivations for completeness.

With either approach, we will be using the Canonical Black-Litterman Reference Model. In this reference model the estimation model is used to compute the distribution of the estimated returns about the mean return, and then to estimate the distribution of returns about the mean return. This distinction is important in understanding the values used for τ and Ω, and for the computations of the variance of the prior and posterior distributions of returns.

The posterior estimate of the mean generated by the estimation model is more precise than either the prior estimate or the investors views. Note that we would not expect large changes in the variance of the distribution of returns about the mean because our estimate of the mean is more precise. 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 either estimation model 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 about the mean in ½ just because we have a slightly better estimate of the mean. In this case, the mean is the random variable, thus the variance of our posterior corresponds to the variance of the estimated mean around the mean return; not the variance of the distribution of returns about the mean return. In this case, the posterior 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, even though we do not have a better estimate of the the actual distribution of returns around the mean.

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 with our problem 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 (2) from our reference model as a starting point.

Our simple linear model is shown below:

(17) $$x\beta = \pi + u$$

Where

π is the n x 1 vector of equilibrium returns for the assets.

x is the n x n matrix I_n which are the factor loadings for our model.
β 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 Φ 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 β and *u* are independent, and *x* is constant, then we can model the variance of π as:

$$V(\pi)=x'V(\beta)x+V(u)$$

Which can be simplified to:

(18) $$V(\pi)=\Sigma+\Phi$$

Where

Σ is the historical covariance matrix of the returns as used earlier.
Φ is the covariance of the residuals or of the estimate about the actual mean.

This ties back to formula (3) 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.

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.

(19) $$p\beta=q+v$$

Where

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

We can combine the prior and conditional information by writing:

$$\begin{bmatrix}x\\p\end{bmatrix}\hat{\beta}=\begin{bmatrix}\pi\\q\end{bmatrix}+\begin{bmatrix}u\\v\end{bmatrix}$$

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

$$V\left(\begin{bmatrix}u\\v\end{bmatrix}\right)=E\left(\begin{bmatrix}u\\v\end{bmatrix}\begin{bmatrix}u' & v'\end{bmatrix}\right)=\begin{bmatrix}\Phi & 0\\0 & \Omega\end{bmatrix}$$

We can then apply the generalized least squares procedure, which leads to estimating $\hat{\beta}$ as

$$\hat{\beta}=\left[\left[x' \quad p'\right]\begin{bmatrix}\Phi & 0\\ 0 & \Omega\end{bmatrix}^{-1}\begin{bmatrix}x\\ p\end{bmatrix}\right]^{-1}\left[x' \quad p'\right]\begin{bmatrix}\Phi & 0\\ 0 & \Omega\end{bmatrix}^{-1}\begin{bmatrix}\pi\\ q\end{bmatrix}$$

This can be rewritten without the matrix notation as

(20) $$\hat{\beta}=\left[x'\Phi^{-1}x+p'\Omega^{-1}p\right]^{-1}\left[x'\Phi^{-1}\pi+p'\Omega^{-1}q\right]$$

For the Black-Litterman model which is a single factor per asset, we can drop x as it is the identity matrix. If one wanted to use a multi-factor model for the equilibrium, then x would be the equilibrium factor loading matrix.

(21) $$\hat{\beta}=\left[\Phi^{-1}+p'\Omega^{-1}p\right]^{-1}\left[\Phi^{-1}\pi+p'\Omega^{-1}q\right]$$

This new $\hat{\beta}$ is the weighted average of the estimates, where the weighting factor is the precision of the estimates. The precision is the inverse of the variance. The posterior estimate $\hat{\beta}$ 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.

If we were using a factor model for the prior, then we would keep x, the factor weightings, in the formulas. This would give us a multi-factor model, where all the factors will be priced into the equilibrium.

We can reformulate our combined relationship in terms of our estimate of $\hat{\beta}$ and a new residual $\tilde{u}$ as

$$\begin{bmatrix}x\\ p\end{bmatrix}\hat{\beta}=\begin{bmatrix}\pi\\ q\end{bmatrix}+\tilde{u}$$

Once again $E(\tilde{u})=0$, so we can derive the expression for the variance of the new residual as:

(22) $$V(\tilde{u})=E(\tilde{u}'\tilde{u})=\left[\Phi^{-1}+p'\Omega^{-1}p\right]^{-1}$$

and the total variance is

$$V(\left[y \quad \pi\right])=V(\hat{\beta})+V(\tilde{u})$$

We began this section by asserting that the variance of the return process is a known quantity. Improved estimation of the quantity $\hat{\beta}$ does not change our estimate of the variance of the return distribution, Σ. 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 (18) to

(23) $$V(\left[y \quad \pi\right])=\Sigma+V(\tilde{u})$$

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.

This is the convention for computing the covariance of the posterior distribution of the canonical reference model as shown in He and Litterman (1999).

In the absence of views,formula (23) simplifies to

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

Which is the variance of the prior distribution of returns.

Appendix A contains a more detailed derivation of formulas (21) and (22).

Using Bayes Theorem for the Estimation Model

See Appendix B for a brief introduction to Bayes Theorem as it is used in deriving the Black-Litterman model.

In the Black-Litterman model, the prior distribution is based on the equilibrium implied excess returns. One of the major assumptions made by the Black-Litterman model is that the covariance of the prior estimate is proportional to the covariance of the actual returns, but the two quantities are independent. The parameter τ will serve as the constant of proportionality. The prior distribution for the Black-Litterman model was specified in formula (11).

The conditional distribution is based on the investor's views. The investor's views are specified as returns to portfolios of assets, and each view has an uncertainty which will impact the overall mixing process. The conditional distribution from the investor's views was specified in formula (13).

The posterior distribution from Bayes Theorem is the precision weighted average of the prior estimate and the conditional estimate. 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 (11) and (13), for our prior and conditional distribution respectively, we can apply Bayes Theorem and derive the following formula for the posterior distribution of asset returns.

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

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

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

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

The derivation of formula (25) 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.

Computing the posterior covariance of returns requires adding the variance of the estimate about the mean, to the variance of the distribution about the estimate the same as in (23). This is mentioned in He and Litterman (1999) but not in any of the other papers.

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

Substituting the posterior variance from (26) we get

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

In the absence of views this reduces to

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

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.[4] Idzorek (2005) and most other authors do not compute a new posterior variance, but instead use the known input variance of the returns about the mean.

In the event our investor has only partial views, that is views on a subset of the 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 τ, but it will be measurable.

Since we often build the known covariance matrix of returns, Σ, from historical data we can use methods from basic statistics to compute τ. as τΣ is analogous to the standard error. 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. Black and Litterman (1992), He and Litterman (1999) and Idzorek (2005) all 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 has no intuitive connection to the data and in fact shows that their paper uses the Alternative Refernce Model.

We can check our results by seeing if the results match our intuition at the boundary conditions.

Given formula (25) it is easy to let $\Omega \rightarrow 0$ showing that the return under 100% certainty of the views is

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

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

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

$$\hat{\Pi} = P^{-1} Q$$

If the investor is not sure about their views, so $\Omega \rightarrow \infty$, then formula (25) reduces to

$$\hat{\Pi} = \Pi$$

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

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

The alternate formula for the posterior variance derived from (26) using the Woodbury Matrix Identity is

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

If $\Omega \rightarrow 0$ (total confidence in views, and every asset is in at least one view) then formula (30) can be reduced to M = 0. If on the other hand the investor is not confident in their views, $\Omega \rightarrow \infty$, then formula (30) can be reduced to M = $\tau\Sigma$.

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), and in the work of Meucci prior to his introduction of "Beyond Black-Litterman".

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

In this reference model, μ is normally distributed with variance Σ. We estimate μ, but μ is not considered a random variable. This is commonly described as having a τ = 1, but more precisely we are making a point estimate and thus have eliminated τ as a parameter. In this model Ω becomes the covariance of the returns to the views around the unknown mean return, , just as Σ is the covariance of the prior return about it's mean. Gven that we are now using point estimates, the posterior is now a point estimate and we no longer worry about posterior covariance of the estimate. In this model we do not have a posterior precision to use downstream in our portfolio selection model.

We rewrite formula (25), noting that we can move around the τ term.

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

Now we see that τ only appears in one term in this formula. Because the Alternative Reference Model does not include updating the Covariance of the estimates this is the only formula. Given that the investor is selecting both Ω and τ to control the blending of the prior and their views, we can eliminate one of the terms. Since τ is a single term for all views and Ω has a separate element for each view, we will keep Ω. We can then rewrite the posterior estimate of the mean, as follows.

(31) $$\hat{\Pi} = \Pi + \Sigma P^T [(P\Sigma P^T) + \Omega]^{-1} [Q - P\Pi]$$

The primary artifacts of this new reference model are, first τ is gone, and second, the investor's portfolio weights in the absence of views equal the equilibrium portfolio weights. Finally at implementation time there is no need or use of formulas (26) or (27).

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 Canonical Reference Model, or of the fact that different authors actually use quite different reference models.

In the Canonical Reference Model, the updated posterior covariance of the unknown mean about the estimate will be smaller than the covariance of either the prior or conditional estimates, indicating that the addition of more information will reduce the uncertainty of the model. The variance of the returns from formula (27) 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 (27) 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. We know that investors using the Canonical Reference Model use τ and that it has a very precise meaning in the model. An author who selects an essentially random value for τ is likely not using the Canonical Reference Model, but is instead using the Alternative Reference Model.

Given the Canonical 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 similar to the one used by He and Litterman (1999). Rather than using only the diagonal, we will retain the entire structure of the covariance matrix to make the math simpler.

(32) $\Omega = P(\tau\Sigma)P^T$

We can substitute this into formula (25) as

$$\begin{aligned}\hat{\Pi} &= \Pi + \tau\Sigma P^T[(P\tau\Sigma P^T)+\Omega]^{-1}[Q-P\Pi^T]\\ &= \Pi + \tau\Sigma P^T[(P\tau\Sigma P^T)+(P\tau\Sigma P^T)]^{-1}[Q-P\Pi^T]\\ &= \Pi + \tau\Sigma P^T[2(P\tau\Sigma P^T)]^{-1}[Q-P\Pi^T]\\ &= \Pi + (\frac{1}{2})\tau\Sigma P^T(P^T)^{-1}[P\tau\Sigma]^{-1}[Q-P\Pi^T]\\ &= \Pi + (\frac{1}{2})\tau\Sigma(\tau\Sigma)^{-1}P^{-1}[Q-P\Pi^T]\\ &= \Pi + (\frac{1}{2})P^{-1}[Q-P\Pi^T]\end{aligned}$$

(33) $\hat{\Pi} = \Pi + (\frac{1}{2})[P^{-1}Q-\Pi^T]$

Clearly using formula (32) 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 $\hat{\Pi}$. In the Canonical Reference Model it does not eliminate τ from the formulas for posterior covariance (27). In the general form if we formulate Ω as

(34) $\Omega = P(\alpha\tau\Sigma)P^T$

Then we can rewrite formula (33) as

(35) $$\hat{\Pi} = \Pi + (\frac{1}{1+\alpha})[P^{-1}Q - \Pi]$$

We can see a similar result if we substitute formula (32) into formula (30).

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

(36) $$M = (\frac{1}{2})\tau\Sigma$$

Note that τ is not eliminated from formula (36). We can also observe that if τ is on the order of 1 and we were to use formula (27) 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 $\alpha \geq 0$, substituting into (25) and following the same logic as used to derive (36) we get

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

This parametrization 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 parametrization of the uncertainty of the views does not have a negative impact on the stability of the results.

Note that this specification of the uncertainty in the views changes the assumption from the views being uncorrelated, to the views having the same correlations as the prior returns.

In summary, if the investor uses the Alternative Reference Model and makes Ω proportional to Σ, then they need only calibrate the constant of proportionality, α, which indicates their relative confidence in their views versus the equilibrium. If they use the Canonical Referemce Model and set Ω proportional to $\tau\Sigma$, then the return estimate will not depend on the value of τ, but the posterior covariance of returns will depend on the proper calibration of τ.

Calibrating τ

This section will discuss some empirical ways to select and calibrate the value of τ.

The first method to calibrate τ 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 inverse of the number of samples. Given that we are estimating the covariance matrix from historical data, then

$$\tau = \frac{1}{T}$$ The maximum likelihood estimator

$$\tau = \frac{1}{T-k}$$ The best quadratic unbiased estimator

T The number of samples

k The number of assets

There are other estimators, but usually, the first definition above is used. Given that we usually aim for a number of samples around 60 (5 years of monthly samples) then τ is on the order of 0.02. This is consistent with several of the papers which indicate they used values of τ on the range (0.025, 0.05).

The most intuitively easy way to calibrate τ is as part of a confidence interval for the prior mean estimates. We will illustrate this concept with a simple example. Consider the scenario where τ = 0.05 and we consider only a single asset with a prior estimate of 7% as the excess return and 15% as the known standard deviation of returns about the mean. For our view we use a confidence interval of (1%, 5%) with 68% confidence.

We keep the ratio of View Precision to Prior Precision fixed between the two scenarios, one with τ = 0.05 and one with τ = 1. In the first scenario even though we use a seemingly small τ = 0.05, our prior estimate has relatively low precision based on the width of the confidence interval, and thus the posterior estimate will be heavily weighted towards the view. In the second scenario with τ = 1, the prior confidence interval is so wide as to make the prior estimate close to worthless. In order to keep the posterior estimate the same across scenarios, the view estimate also has a wide confidence interval indicating the investor is really not confident in any of their estimates.

τ	Prior @ 68% Confidence	Prior Precision	View σ	View Precision	View @ 68% Confidence	View/Prior Precision
0.05	(4.6%, 9.4%)	888	2.00%	2500	(1%, 5%)	2.81
1.00	(-8%, 22%)	44.4	8.90%	125	(-5.9%, 11.9%)	2.81

Given such wide intervals for the τ = 1 scenario, 16% confidence that our asset has a mean return less than -8%, it is hard to imagine having much conviction in using the final asset allocation. Understanding the interplay between the selection of τ and the specification of the variance of the views is critical.

Note that this example illustrates the difference between the parameters for the Canonical Reference

Model and the Alternative Reference Model. Specifying $\tau = 1$ is the Alternative Reference Model, but just generates garbage outputs from the Canonical Reference Model.

Finally, we could instead calibrate τ 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

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

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.

Results

This section of the document will step through a comparison of 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. Selected formulas are also available as MATLAB and/or SciLab scripts on the website blacklitterman.org. 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 Π 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 (10).

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

As a first test of our algorithm we verify that when the investor has no views that the weights are correct, substituting formula (28) into (10) we get

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

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

Given this result, it is clear that the output weights with no views will be impacted by the choice of τ 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 (38).

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 as they clearly implement the Canonical Reference Model and they provide all the data required to reproduce their results in the paper.

He and Litterman (1999) set

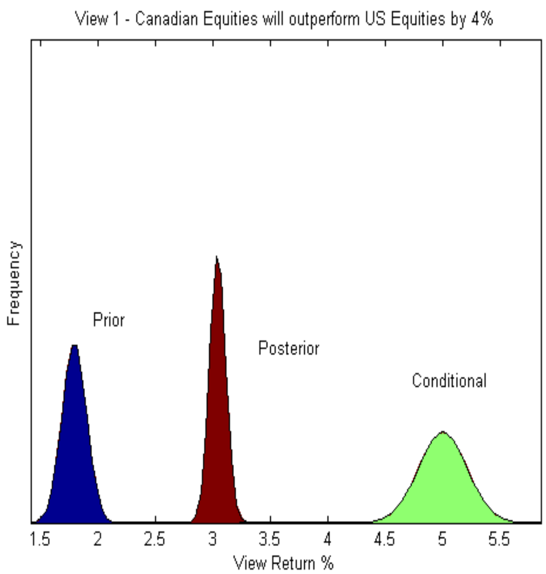
(39) $\Omega = \text{diag}(P^T(\tau\Sigma)P)$

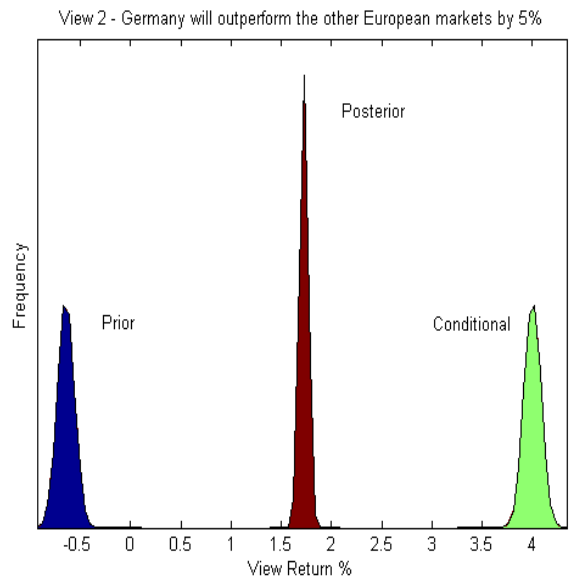
This essentially makes the uncertainty of the views equivalent to the uncertainty of the equilibrium estimates. They select a small value for τ (0.05), and they use the Canonical Reference Model and the updated posterior variance of returns as calculated in formulas (30) and (27).

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.Figure 5 - Distributions of Actual Means about Estimates Means





The preceding diagrams show the pdf for the prior, view and posterior for each View defined in the problem. The y axis uses the same scale in each graph. Note how in view 1 the view (conditional distribution of the estimated mean) is much more diffuse because the variance of the estimate is larger (precision of the estimate is smaller). Also note how the precision of the prior and views impacts the precision (width of the peak) on the pdf. In view 1 with the less precise view, the posterior is also less precise.

Matching the Results of Idzorek

This section of the document describes the efforts to reproduce the results of Idzorek (2005). In trying to match Idzorek's results we 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 Canonical Reference Model,, but in the end the differences amounted to only 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 (2005) and the Canonical Reference Model.. Table 3 shows the same results as generated by the Alternative Reference Model.

Table 2 – Canonical Reference Model with Idzorek data (Corresponds with data in Idzorek's Table 6).

Asset Class	μ	w_{eq}	w*	Black-Litterman Reference Model	Idzorek's Results
US Bonds	.07	18.87%	28.96%	10.09%	10.54
Intl Bonds	.50	25.49%	15.41%	-10.09%	-10.54

Asset Class	μ	w_{eq}	w*	Black-Litterman Reference Model	Idzorek's Results
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 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 with Idzorek data (Corresponds with data in Idzorek's Table 6).

Country	μ	w_{eq}	w	Alternative Reference Model	Idzorek's Results
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

Additional Work

This section provides a brief discussion of efforts to reproduce results from some of the major research papers on the Black-Litterman model.

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 Alernative 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 we have been unable to formulate as of this time. We plan on continuing this work with the goal of verifying the details of the Black-Litterman implementation used in Black and Litterman (1992).

Extensions to the Black-Litterman Model

In this section we will cover the extensions to the Black-Litterman model proposed in Idzorek (2005), Fusai and Meucci (2003), Krishnan and Mains (2006) and Qian and Gorman (2001).

Idzorek (2005) presents a means to calibrate the confidence or variance of the investors views in a simple and straightforward method.

Next is a section on measures of extremity or quality of views. 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. Braga and Natale (2007) describe how to use Tracking Error to measure the distance from the equilibirum to the posterior portfolio. I also include additional original work on using relative entropy to measure quality of the views.

Finally, larger extensions to the model such as Krishnan and Mains (2006) present a method to incorporate additional factors into the model. Qian and Gorman (2001) present a method to integrate views on the covariance matrix as well as views on the returns.

Idzorek's Extension

Idzorek's apparent goal was to reduce the complexity of the Black-Litterman model for non-quantitative investors. He achieves this by allowing the investor to specify the investors confidence in the views 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

(40) $$confidence = (\hat{w} - w_{mkt}) / (w_{100} - w_{mkt})$$

w_{100} is the weight of the asset under 100% certainty in the view

w_{mkt} is the weight of the asset under no views

w is the weight of the asset under the specified view.

He provides a method to back out the value of ω required to generate the proper tilt (change in weights from prior to posterior) for each view. These values are then combined to form Ω, and the model is used to compute posterior estimates.

Idzorek includes τ in his formulas, but because of his use of the Alternative Reference Model and his formula (40) there is no need to use τ with the Idzorek method.

In his paper he discusses solving for ω using a least squares method. We can actually solve this analytically[5]. The next section will provide a derivation of the formulas required for this solution.

First we will use the following form of the uncertainty of the views. Idzorek includes τ in this formula, but he uses the Alternative Reference Model so we can drop τ from the formulation of his method.

(41) $$\Omega = \alpha P \Sigma P^T$$

α, the coefficient of uncertainty, is a scalar quantity in the interval [0, ∞]. When the investor is 100% confident in their views, then α will be 0, and when they are totally uncertain then α will be ∞. Note that formula (41) is exact, it is identical to formula (39) the Ω used by He and Litterman (1999) because it is a 1x1 matrix. This allows us to find a closed form solution to the problem of Ω for Idzorek's confidence.

First we substitute formula (37)

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

into formula (10).

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

Which yields

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

Now we can solve formula (42) at the boundary conditions for α.

$$\lim_{\alpha\to\infty}, w_{mk} = \Pi(\delta\Sigma)^{-1}$$

$$\lim_{\alpha\to 0}, w_{100} = P^{-1}Q(\delta\Sigma)^{-1}$$

And recombining some of the terms in (42) we arrive at

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

Substituting w_{mk} and w_{100} back into (43) we get

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

5 Thanks to Boris Gnedenko for pointing this fact out to me.

And comparing the above with formula (40)

$$confidence = (\hat{w} - w_{mk}) / (w_{100} - w_{mk})$$

We see that

$$confidence = \frac{1}{(1+\alpha)}$$

And if we solve for α

(44) $\alpha = (1 - confidence) / confidence$

Using formulas (44) and (41) the investor can easily calculate the value of ω for each view, and then roll them up into a single Ω matrix. To check the results for each view, we then solve for the posterior estimated returns using formula (25) and plug them back into formula (40). Note that when the investor applies all their views at once, the interaction amongst the views may pull the posterior estimate for individual assets away from the results generated when the views were taken one at a time.

Idzorek's method greatly simplifies the investor's process of specifying the uncertainty in the views when the investor does not have a quantitative model driving the process. In addition, this model does not add meaningful complexity to the process.

An Example of Idzorek's Extension

Idzorek describes the steps required to implement his extension in his paper, but does not provide a complete worked example. In this section I will work through his example from where he leaves off in the 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 his paper Idzorek defines the steps in his method which include calculations of w_{100} and then the calculation of ω for each view given the desired change in the weights. From the previous section, we can see that we only need to take the investor's confidence for each view, plug it into formula (44) and compute the value of alpha. Then we plug α, P and Σ into formula (41) and compute the value of ω for each view. At this point we can assemble our Ω matrix and proceed to solve for the posterior returns using formula (24) or (25).

In working the example, I will show the results for each view including the w_{mkt} and w_{100} in order to make the workings of the extension more transparent. Tables 4, 5 and 6 below each show the results for a single view.

Table 4 – Calibrated Results for View 1

Asset	ω	w_{mkt}	w*	$w_{100\%}$	Implied Confidence
Intl Dev Equity	.002126625	24.18%	25.46%	29.28%	25.00%

Table 5 – Calibrated Results for View 2

Asset	ω	w_{mkt}	w*	$w_{100\%}$	Implied Confidence
US Bonds	.000140650	19.34%	29.06%	38.78%	50.00%
Intl Bonds	.000140650	26.13%	16.41%	6.69%	50.00%

Table 6 – Calibrated Results for View 3

Asset	ω	w_{mkt}	w*	$w_{100\%}$	Implied Confidence
US LG	.000466108	12.09%	9.49%	8.09%	65.00%
US LV	.000466108	12.09%	14.69%	16.09%	65.00%
US SG	.000466108	1.34%	1.05%	.90%	65.00%
US SV	.000466108	1.34%	1.63%	1.78%	65.00%

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%

Asset	View 1	View 2	View 3	μ	σ	w_{mkt}	Posterior Weight	change
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	.08507	.00563	.01864					
Lambda	.002	-.006	-.002					

Then we use the freshly computed values for the Ω matrix with all views specified together and arrive at the final result shown above in Table 7 blending all 3 views together.

Measuring the Impact of the Views

This section will discuss several methods used in the literature to measure the impact of the views on the posterior distribution. In general we can divide these measures into two groups. The first group allows us to test the hypothesis that the views or posterior contradict the prior. The second group allows us to measure a distance or information content between the prior and posterior.

Theil (1971), and Fusai and Meucci (2003) describe measures that are designed to allow a hypothesis test to ensure the views or the posterior does not contradict the prior estimates. Theil (1971) describes a method of performing a hypothesis test to verify that the views are compatible with the prior. We will extend that work to measure compatibility of the posterior and the prior. Fusai and Meucci (2003) describe a method for testing the compatibility of the posterior and prior when using the alternative reference model.

He and Litterman (1999), and Braga and Natale (2007) describe measures which can be used to measure the distance between two distributions, or the amount of tilt between the prior and the posterior. These measures don't lend themselves to hypothesis testing, but they can potentially be used as constraints on the optimization process. He and Litterman (1999) define a metric, Λ, which measures the tilt induced in the posterior by each view. Braga and Natale (2007) use Tracking Error Volatility (TEV) to measure the distance from the prior to the posterior.

Theil's Measure of Compatibility Between the Views and the Prior

Theil (1971) describes this as testing the compatibility of the views with the prior information. Given the linear mixed estimation model, we have the prior (17) and the conditional (19).

(17) $x\beta=\pi+u$

(19) $p\beta = q + v$

The mixed estimation model defines *u* as a random vector with mean 0 and covariance $\tau\Sigma$, and v as a random vector with mean 0 and covariance Ω.

The approach we will take is very similar to the approach taken when analyzing a prediction from a linear regression. We have two estimates of the views, the conditional estimate of the view and the posterior estimate of the returns.

We can define the posterior estimate as $\hat{\beta}$. We can measure the estimation error between the prior and the views as:

(45) $\zeta = (x\hat{\beta} - \mu) = x(\hat{\beta} - \beta) + u$

The vector ζ has mean 0 and variance $V(\zeta)$. We will form our hypothesis test using the formulation

(46) $\xi = E(\zeta) V(\zeta)^{-1} E(\zeta)$

The quantity, ξ, is known as the Mahalanobis distance (multi-dimensional analog of the z-score) and is distributed as $\chi^2(n)$. In order to use this form we need to solve for the $E(\zeta)$ and $V(\zeta)$.

If we consider only the information in the views, the estimator of β is:

$$\hat{\beta} = (P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} Q$$

Note that since P is not required to be a matrix of full rank that we might not be able to evaluate this formula as written. We work in return space here (as opposed to view space) as it seems more natural. Later on we will transform the formula into view space to make it computable.

We then substitute the new estimator into the formula (45) and eliminate x as it is the identity matrix in the Black-Litterman application of mixed-estimation.

(47) $\zeta = (P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} Q - \beta + u$

Next we substitute formula (19) for Q.

$$\zeta = (P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} (P\beta + v) - \beta + u$$
$$\zeta = (P^T \Omega^{-1} P)^{-1} (P^T \Omega^{-1} P)\beta + (P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} v - \beta + u$$
$$\zeta = (P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} v + u$$

Give our estimator, we want to find the variance of the estimator.

$$V(\zeta) = E(\zeta \zeta^T)$$
$$V(\zeta) = E((P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} v v^T \Omega^{-1} P (P^T \Omega^{-1} P)^{-1} - 2(P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} v u^T + u u^T)$$

But *E*(vu) = 0 so we can eliminate the cross term, and simplify the formula.

$$V(\zeta) = E[(P^T \Omega^{-1} P)^{-1} P^T \Omega^{-1} v v^T \Omega^{-1} P (P^T \Omega^{-1} P)^{-1} + u u^T]$$
$$V(\zeta) = E[(P^{-1} v v^T (P^T)^{-1}) + u u^T]$$
$$V(\zeta) = (P^T \Omega^{-1} P)^{-1} + \tau \Sigma$$

The last step is to take the expectation of formula (47). At the same time we will substitute the posterior estimate (μ) for β.

$$E(\zeta)=(P^T\Omega^{-1}P)^{-1}P^T\Omega^{-1}Q-\Pi$$

Now substitute the various values into (46) as follows

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

Unfortunately, under the usual circumstances we cannot compute ξ. Because P does not need to contain a view on every asset, several of the terms are not always computable as written. However, we can easily convert it to view space by multiplying by P and P^T.

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

This new test statistic statistic, ξ, in formula (48) is distributed as $\chi^2(q)$ where q is the number of views. It is the square of the Mahalanobis distance of the posterior return estimate versus the posterior covariance of the estimate. We can use this test statistic to determine if our views are consistent with the prior by means of a standard confidence test.

$$P(q)=1-F(\xi(q))$$

Where F(ξ) is the CDF of the $\chi^2(q)$ distribution.

We can also compute the sensitivities of this measure to the views using the chain rule.

$$\frac{\partial P}{\partial q}=\frac{\partial P}{\partial \xi}\frac{\partial \xi}{\partial q}$$

Substituting the various terms

(49) $$\frac{\partial P}{\partial q}=-f(\xi)\left[2((\Omega+P\tau\Sigma P^T)^{-1}(Q-P\Pi)^T)\right]$$

Where f(ξ) is the PDF of the $\chi^2(q)$ distribution.

Theil's Measure of the Source of Posterior Information

Theil (1963) describes a measure which can be used to determine the contribution to the posterior precision of the prior and the views. This measure which he calls θ sums to 1 across all sources, and conveniently also sums across the views if we measure the contribution of each view.

The measure for the contribution to posterior precision from the prior is

(50) $$\theta_{prior}=\frac{1}{n}tr((\tau\Sigma)^{-1}[(\tau\Sigma)^{-1}+P^T\Omega^{-1}P]^{-1})$$

where n is the number of assets. The formula below can be used for all views by using the full matrices P and Ω. For a single view (i) just use the relevant slices of P and Ω.

(51) $$\theta_i=\frac{1}{n}tr((P_i^T\Omega_{i,i}^{-1}P_i)[(\tau\Sigma)^{-1}+P^T\Omega^{-1}P]^{-1})$$

Formulas (50) and (51) provide the formulas we can use to compute the contribution towards the posterior precision from the prior and from the views. We can use this diagnostic to identify if the relative contributions match our intuitive view on this proportion.

Fusai and Meucci's Measure of Consistency

Next we will look at the work 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.

In their paper they use the Alternative Reference Model. Their measure is analagous to Theil's Measure of Compatibility, but because the alternative reference model uses the prior variance of returns for the posterior they do not need any derivation of the variance. We can apply a variant of their measure to the Canonical Reference Model as well.

They propose the use of the squared Mahalanobis distance of the posterior returns from the prior returns. We include τ here to match the Canonical Reference Model, but their work does not include τ as they use the Alternative Reference Model.

(52) $$M(q)=(\mu_{BL}-\mu)(\tau\Sigma)^{-1}(\mu_{BL}-\mu)$$

It is essentially measuring the distance from the prior, μ, to the estimated returns, μ_{BL}, normalized by the uncertainty in the estimate. We use the covariance matrix of the prior distribution as the uncertainty. The squared Mahalanobis distance is distributed as $\chi^2(q)$ where q is the number of assets.. This can easily be used in a hypothesis test. Thus the probability of this event occurring can be computed as:

(53) $$P(q)=1-F(M(q))$$

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

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

(54) $$\frac{\partial P(q)}{\partial q}=-f(M)\left[2(\mu_{BL}-\mu)\right]\left[(P(\tau\Sigma)P+\Omega)^{-1}P\right]$$

Where f(M) is the PDF of the chi square distribution with n degrees of freedom for M(q). Note that this measure is very similar to (49) Theil's measure of compatibility between the prior and the views.

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% (a commonly used confidence level in hypothesis testing), 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 providing the largest marginal increase in their measure and they investor can then adjust these views. These sensitivities are especially useful since some views may actually be pulling the posterior towards the prior, and the investor could strengthen these views, or weaken views which pull the posterior away from the prior. This last point may seem non-intuitive. Given that the views are indirectly coupled by the covariance matrix, one would expect that the views only push the posterior distribution away from the prior. However, because the views can be conflicting, either directly or via the correlations, any individual view can have a net impact pushing the posterior closer to the prior, or pushing it further away.

They propose to use their measure in an iterative method to ensure that the posterior is consistent with the prior to the specified confidence level.

With the Canonical Reference Model we could rewrite formula (52) using the posterior variance of the return instead of $\tau\Sigma$ yielding:

(55) $$M(q)=(\mu_{BL}-\mu)((\tau\Sigma)^{-1}+P^T\Omega^{-1}P)^{-1}(\mu_{BL}-\mu)$$

Otherwise their Consistency Measure and it's use is the same for both reference models.

He and Litterman Lambda

He and Litterman (1999) use a measure, Λ, to measure the impact of each view on the posterior. They define the Black-Litterman unconstrained posterior portfolio as a blend of the equilibrium portfolio (prior) and a contribution from each view, that contribution is measured by Λ.

Deriving the formula for Λ we will start with formula (10) and substitute in the various values from the posterior distribution.

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

We substitute the return from formula (0) for $\hat{\Pi}$

(56) $$\hat{w}=\frac{1}{\delta}\tilde{\Sigma}^{-1}M^{-1}[(\tau\Sigma)^{-1}\Pi+P^T\Omega^{-1}Q]$$

We will first simplify the covariance term

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

Then we can define

(57) $$A = \left[\frac{\Omega}{\tau} + \frac{P^T \Sigma P}{1+\tau}\right]$$

And finally rewrite as

(58) $$\tilde{\Sigma}^{-1} M^{-1} = \frac{\tau}{1+\tau} [I - (P^T A^{-1} P)(\frac{\Sigma}{1+\tau})]$$

Our goal is to simplify the formula to the form

$$\hat{w} = \frac{1}{1+\tau} \left\{ w_{eq} + P^T \Lambda \right\}$$

We use the multiplier, $\frac{1}{1+\tau}$, because in the Black-Litterman Reference Model the investor is not fully invested in the prior (equilibrium) portfolio.

In order to find Λ we substitute formula (58) into (56) and then gather terms.

$$\hat{w} = \frac{1}{\delta} \frac{\tau}{1+\tau} \left[I - (P^T A^{-1} P)(\frac{\Sigma}{1+\tau}) \right] \left[(\tau \Sigma)^{-1} \Pi + P^T \Omega^{-1} Q \right]$$
$$\hat{w} = \frac{1}{\delta} \frac{\tau}{1+\tau} \left[(\tau \Sigma)^{-1} \Pi - (P^T A^{-1} P)(\frac{\tau}{1+\tau}) \Pi + P^T \Omega^{-1} Q - (P^T A^{-1} P)(\frac{\Sigma}{1+\tau}) P^T \Omega^{-1} Q \right]$$
$$\hat{w} = \frac{1}{1+\tau} \left\{ w_{eq} + P^T \left[-A^{-1} P \Pi (\frac{1}{\delta(1+\tau)}) + \frac{\tau \Omega^{-1} Q}{\delta} - (A^{-1} P)(\frac{\Sigma}{\delta(1+\tau)}) P^T \Omega^{-1} Q \right] \right\}$$
$$\hat{w} = \frac{1}{1+\tau} \left\{ w_{eq} + P^T \left[\frac{\tau}{\delta} \Omega^{-1} Q - \frac{A^{-1} P \Sigma w_{eq}}{1+\tau} - A^{-1} \frac{\tau}{1+\tau} (P^T \Sigma P) \frac{\Omega^{-1} Q}{\delta} \right] \right\}$$

So we can see that along with (57), the following formula defines Λ.

(59) $$\Lambda = \frac{\tau}{\delta}\Omega^{-1}Q - \frac{A^{-1}P\Sigma w_{eq}}{1+\tau} - A^{-1}\frac{\tau}{1+\tau}(P\Sigma P^T)\frac{\Omega^{-1}Q}{\delta}$$

He and Litterman Λ represents the weight on each of the view portfolios on the final posterior weight. As a result, we can use Λ as a measure of the impact of our views.

We can also derive the He and Litterman Λ for the Alternative Reference Model, we will call this Λ_A. We start from the same reverse optimization formula.

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

We substitute the return from formula (24) for $\hat{\Pi}$, but we use the prior covariance matrix as we are using the Alternative Reference Model.

(60) $$\hat{w} = \frac{1}{\delta}\Sigma^{-1}((\Sigma)^{-1} + P^T\Omega^{-1}P)^{-1}[\Sigma^{-1}\Pi + P^T\Omega^{-1}Q]$$

We will first simplify the covariance term using the Woodbury Matrix Identity.

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

Then we can define

$$A_A = \Omega + P^T\Sigma P$$

Then we can substitute the above result back into formula (43) and expand the terms.

$$\hat{w} = \frac{1}{\delta}(I - P^T A_A^{-1}P\Sigma)[\Sigma^{-1}\Pi + P^T\Omega^{-1}Q]$$
$$\hat{w} = \frac{1}{\delta}(\Sigma^{-1}\Pi + P^T\Omega^{-1}Q - P^T A_A^{-1}P\Pi - P^T A_A^{-1}P\Sigma P^T\Omega^{-1}Q)$$
$$\hat{w} = \frac{\Sigma^{-1}\Pi}{\delta} + \frac{1}{\delta}P^T\left[(\Omega^{-1}Q - A_A^{-1}P\Pi - A_A^{-1}P\Sigma P^T\Omega^{-1}Q)\right]$$
$$\hat{w} = \frac{\Sigma^{-1}\Pi}{\delta} + \frac{1}{\delta}P^T\left[(\Omega^{-1}Q - A_A^{-1}P\delta\Sigma w_{eq} - A_A^{-1}P\Sigma P^T\Omega^{-1}Q)\right]$$

He and Litterman Λ takes the similar form in the Alternative Reference Model as shown below.

$$\hat{w} = w_{eq} + P^T\Lambda_A$$

Note that in the Alternative Reference Model, the investor's prior portfolio has the exact same weights as the equilibrium portfolio.

We can see that the following formula defines Λ_A.

(61) $$\Lambda_A = \frac{\Omega^{-1} Q}{\delta} - A_A^{-1} P \Sigma w_{eq} - A_A^{-1} P \Sigma P^T \frac{\Omega^{-1} Q}{\delta}$$

Braga and Natale and Tracking Error Volatility

Braga and Natale (2007) propose the use of tracking error between the posterior and prior portfolios as a measure of distance from the prior. Tracking error is commonly used by investors to measure risk versus a benchmark, and can be used as an investment constraint. As it is so commonly used, most investors have an intuitive understanding and a level of comfort with TEV. Tracking error volatility is defined as

(62) $$TEV = \sqrt{w_{actv}^T \Sigma w_{actv}} \quad where\ w_{actv} = \hat{w} - w_r$$

Where

w_{actv}	Active weights, or active portfolio
$\hat{w}$	Weight in the investor's portfolio
w_r	Weight in the reference portfolio
Σ	Covariance matrix of returns

They also derive the formula for tracking error sensitivities as follows: Given that

$$TEV = f(w_{actv})$$

and we can further refine

$$w_{actv} = g(q) \quad where\ q\ represents\ the\ views$$

Then we can use the chain rule to decompose the sensitivity of TEV to the views

(63) $$\frac{\partial TEV}{\partial q} = \frac{\partial TEV}{\partial w_{actv}} \frac{\partial w_{actv}}{\partial q}$$

We can solve for the first term of formula (63) directly,

$$\frac{\partial TEV}{\partial w_{actv}} = \partial \frac{(\sqrt{w_{actv}^T \Sigma w_{actv}})}{\partial w_{actv}}$$

$$Let\ x = w_{actv}^T \Sigma w_{actv},\ \ then\ apply\ the\ chain\ rule$$

$$\frac{\partial TEV}{\partial w_{actv}} = \frac{\partial TEV}{\partial x} \frac{\partial x}{\partial w_{actv}}$$

$$\frac{\partial TEV}{\partial w_{actv}} = [\frac{1}{2\sqrt{w_{actv}^T \Sigma w_{actv}}}][2 \Sigma w_{actv}]$$

$$\frac{\partial TEV}{\partial w_{actv}} = \frac{\Sigma w_{actv}}{\sqrt{w_{actv}^T \Sigma w_{actv}}}$$

Solving for the second term of formula (63) is slightly more complicated.

$$\frac{\partial w_{actv}}{\partial q} = \frac{\partial(\hat{w} - w_{ref})}{\partial q}$$
$$\frac{\partial w_{actv}}{\partial q} = \frac{\partial((\delta\Sigma)^{-1}E(r) - (\delta\Sigma)^{-1}\Pi)}{\partial q}$$
$$\frac{\partial w_{actv}}{\partial q} = (\delta\Sigma)^{-1}\frac{\partial(E(r) - \Pi)}{\partial q}$$
$$\frac{\partial w_{actv}}{\partial q} = (\delta\Sigma)^{-1}\partial\frac{\tau\Sigma P^T[(P\tau\Sigma P^T) + \Omega]^{-1}[Q - P\Pi]}{\partial Q}$$
$$\frac{\partial w_{actv}}{\partial q} = (\delta\Sigma)^{-1}\tau\Sigma P^T[(P\tau\Sigma P^T) + \Omega]^{-1}$$
$$\frac{\partial w_{actv}}{\partial q} = \frac{1}{\delta}P^T[(P\Sigma P^T) + \frac{\Omega}{\tau}]^{-1}$$

This result is somewhat different from that found in the Braga and Natale (2007) paper because we use the form of the Black-Litterman model which requires less matrix inversions. The formula for the sensitivities is

(64) $$\frac{\partial TEV}{\partial q} = \frac{\Sigma w_{actv}}{\sqrt{w_{actv}^T \Sigma w_{actv}}}\frac{1}{\delta}P^T\left[(P\Sigma P^T) + \frac{\Omega}{\tau}\right]^{-1}$$

We can come up with the equivalent metric for the Canoni6`1cal Reference Model model quite easily. If we just notice that in a tracking error scenario, the covariance matrix should be the most accurate which would be the posterior covariance matrix.

$$\frac{\partial TEV}{\partial q} = \frac{\tilde{\Sigma} w_{actv}}{\sqrt{w_{actv}^T \tilde{\Sigma} w_{actv}}}\frac{\partial w_{actv}}{\partial q}$$
$$\frac{\partial TEV}{\partial q} = \frac{\tilde{\Sigma} w_{actv}}{\sqrt{w_{actv}^T \tilde{\Sigma} w_{actv}}}\frac{P^T}{1+\tau}\frac{\partial\Lambda}{\partial q}$$

Yielding the TEV sensitivities for the Canonical Reference Model.

(65) $$\frac{\partial TEV}{\partial q} = \frac{\tilde{\Sigma} w_{actv}}{\sqrt{w_{actv}^T \tilde{\Sigma} w_{actv}}}\frac{P^T}{1+\tau}\left[\frac{\tau}{\delta}\Omega^{-1} - A^{-1}\frac{\tau}{1+\tau}(P\Sigma P^T)\frac{\Omega^{-1}}{\delta}\right]$$

Braga and Natale work through a fairly simple example in their paper, but they do not provide all the raw data required to reproduce their results. I have been unable to reproduce either their equilibrium or their mixing results. Given their posterior distribution as presented in the paper one can easily reproduce their TEV results.

One advantage of the TEV is that most investors are familiar with it, and so they will have some intuition as to what it represents. The consistency metric introduced by Fusai and Meucci (2003) will not be as familiar to investors.

Metrics Introduced in Herold (2003)

Herold (2003) discusses the concept of diagnostics which an investor can apply to their model in order to validate the outputs. One of these diagnostics is the correlation between views as one diagnostic which can be used to determine how the updated portfolio may perform. By examining the correlation matrix of the views $P \Sigma P^t$ we can determine how correlated the views are. If the views are highly correlated then we can expect all views to contribute to out performance if they are correct and underperformance if they are incorrect. If the views are not highly correlated then we can expect a diversified contribution to performance.

As an example, if we use He and Litterman's data and their two views we compute this measure as

$$P \Sigma P^t = \begin{bmatrix} 0.0213 & 0.0020 \\ 0.0020 & 0.0170 \end{bmatrix}$$

The off-diagonal elements, 0.0020, are one order of magnitude smaller than the on-diagonal elements, (0.0213 and 0.0170). This difference indicates that the views are not strongly correlated. This is consistent with He and Littermans' specification of the two views on mutually exclusive sets of assets which are loosely correlated.

Herold (2003) also discusses the marginal contribution to tracking error by view. This is the same metric computed by Braga and Natale, though he proposes deriving the formula for it's calculation from the Alternative Reference Model version of He and Litterman's Λ, Λ_A.

Herold (2003) considers the case of Active Management, which means that the Black-Litterman model is being applied to an active management overlay on some portfolio. The prior distribution in this case corresponds to $w_{eq} = 0$. The posterior weights are the overlay weights. He uses the Alternative Reference Model and computes a quantity, Φ, which is the active management version of Λ_A.

$$\Lambda_A = \frac{\Omega^{-1} Q}{\delta} - A_A^{-1} P \Sigma w_{eq} - A_A^{-1} P \Sigma P^T \frac{\Omega^{-1} Q}{\delta}$$

$$\mathit{but}\ w_{eq} = 0, \quad \mathit{so}$$

$$\Phi = \frac{\Omega^{-1} Q}{\delta} - A_A^{-1} P \Sigma P^T \frac{\Omega^{-1} Q}{\delta}$$

$$\mathit{where} \quad w_{actv} = P^T \Phi$$

Herold starts from formula (62) just as Braga and Natale, but then uses an alternative formula for tracking error.

$$TEV = \sqrt{w_{actv}^T \Sigma w_{actv}} \quad \mathit{where}\ w_{actv} = \hat{w} - w_r$$

and also

$$w_{actv} = \hat{w} - w_{eq} = P^T \Phi$$

$$TEV = \sqrt{(\Phi^T P \Sigma P^T \Phi)}$$

Then we can just take $\frac{\delta\, TEV}{\delta\, Q}$ to find the marginal contribution to tracking error by views. By the chain rule

$$\frac{\delta\, TEV}{\delta\, Q} = \frac{\delta\, TEV}{\delta\, \Phi} \times \frac{\delta\, \Phi}{\delta\, Q}$$

$$\frac{\delta\, TEV}{\delta\, Q} = \frac{P \Sigma P^T \Phi}{\sqrt{(\Phi P \Sigma P^T \Phi)}} \times \frac{1}{\delta}\left[\Omega^{-1} - (P \Sigma P^T + \Omega)^{-1} P \Sigma P^T \Omega^{-1}\right]$$

Note that we could perform the same calculation for the Canonical Reference Model using He and Littermans' Λ, but we would also need to use the posterior covariance of the distribution rather than just Σ in the calculations.

A Demonstration of the Measures

Now we will work a sample problem to illustrate all of the metrics, and to provide some comparison of their features. We will start with the equilibrium from He and Litterman (1999) and for Example 1 use the views from their paper, Germany will outperform other European markets by 5% and Canada will outperform the US by 4%.

Table 8 – Example 1 Returns and Weights, equilibrium from He and Litterman, (1999).

Asset	P_0	P_1	μ	μ_{eq}	$w_{eq}/(1+\tau)$	w*	$w^* - w_{eq}/(1+\tau)$
Australia	0.0	0.0	4.45	3.9	1.50%	1.5%	0.0%
Canada	0.0	1.0	9.06	6.9	2.1%	53.3%	51.2%
France	-0.295	0.0	9.53	8.4	5.0%	-3.3%	-8.3%
Germany	1.0	0.0	11.3	9	5.2%	33.1%	27.9%
Japan	0.0	0.0	4.65	4.3	11.0%	11.0%	0.0%
UK	-0.705	0.0	6.98	6.8	11.8%	-7.8%	-19.6%
USA	0.0	-1.0	7.31	7.6	58.6%	7.3%	-51.3%
q	5.0	4.0					
ω/τ	0.02	.017					

Table (8) illustrates the results of applying the views and Table (0) displays the various impact measures.

Table 9 - Impact Measures for Example 1

Measure	Value (Confidence Level)	Sensitivity (V1)	Sensitivity (V2)
Theil's Measure	1.672 (43.3%)	-5.988	-11.46
Theil's θ	0.858 (prior)	0.0712	0.0712
Fusai and Meucci's Measure	0.8728 (99.7%)	-0.1838	-0.3327
Λ		0.2920	0.5380
TEV	8.28%	0.688	1.294

If we examine the change in the estimated returns vs the equilibrium, we see where the USA returns decreased by only 29 bps, but the allocation decreased 51.25% caused by the optimizer favoring Canada whose returns increased by 216 bps and whose allocation increased by 51.25%. This shows that what appear to be moderate changes in return forecasts can cause very large swings in the weights of the assets, a common issue with unconstrained mean variance optimization. Here we use unconstrained mean variance optimization not because we have to, but because it is well understood and transparent.

Next looking at the impact measures, Theil's measure indicates that we can be confident only at the 43% level that the views are consistent with the prior. If we examine the diagram (5) we can see that indeed these views are significantly different from the prior. Fusai and Meucci's measure of compatibility of the prior and posterior on the other hand comes with a confidence level of 99.8% so they are much more confident that the posterior is consistent with the prior. A major difference in their approaches is that Theil is working in view space, which for our example has order = 2 whilst Fusai and Meucci are working in asset space which has order = 7.

All of the metrics' sensitivities to the views indicate that the second view has a relative weight just about twice as large as the first view, so the second view contributes much more to the change in the weights.

The TEV of the posterior portfolio is 8.28% which is significant in terms of how closely the posterior portfolio will track the equilibrium portfolio. This seems to be a very large TEV value given the scenario we're working with

Next we change our confidence in the views by multiplying the variance by 2, this will increase the change from the prior to the posterior and allow us to make some judgements based on the impact measures.

Table 10 – Example 2 Returns and Weights, equilibrium from 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	4.72	1.5%	1.5%	0.0%
Canada	0.0	1.0	10.3	2.1%	22.7%	20.6%

Asset	P_0	P_1	μ	$w_{eq}/(1+\tau)$	w*	w* - $w_{eq}/(1+\tau)$
France	-0.295	0.0	10.2	5.0%	1.6%	-3.4%
Germany	1.0	0.0	12.4	5.2%	16.8%	11.6%
Japan	0.0	0.0	4.84	11.0%	11.0%	0.0%
UK	-0.705	0.0	7.09	11.8%	3.7%	-8.1%
USA	0.0	-1.0	7.14	58.6%	38.0%	-20.6%
q	5.0	4.0				
ω/τ	0.09	0.07				

Table 11 - Impact Measures for Example 2

Measure	Value (Confidence Level)	Sensitivity (V1)	Sensitivity (V2)
Theil's Measure	1.537 (46.4%)	-4.27	-12.39
Theils θ	0.88	0.05	0.07
Fusai and Meucci's Measure	0.7479 (99.8%)	-0.06	-0.24
Λ		0.193	0.544
TEV	7.67%	0.332	1.406

Examining the updated results in Table (10) we see that the changes to the forecast returns have decreased consistent with our increased uncertainty in the views. We now have an 20% increase in the allocation to Canada and an 20% decrease in the allocation to the USA. From Table (11) we can see that Theil's measure has increased, but only a fraction and we continue to have little confidence that the views our consistent with the prior. Fusai and Meucci's measure now is 99.8% confident that the posterior is consistent with the prior. It is unclear in practice what bound we would want to use, but their measure usually presents which a much higher confidence than Theil's measure. The TEV has decreased, and is now 7.67% which is not significantly smaller.

Once again all the sensitivities show the second view having more of an impact on the final weights.

Across both scenarios we can draw some conclusions about these various measures. Theil's consistency measure test ranged from a high of 46% confident to a low of 44% confident that the views were consistent with the prior estimates. This measure is very sensitive and it is unclear what would be a good confidence threshold. 50% does seem intuitively appealing.

Fusai and Meucci's Consistency measure ranged from 98.38% to 99.98% confident, indicating that the posterior estimates was generally highly consistent with the prior. Fusai and Meucci present that an investor may have a requirement that the confidence level be 5%. In light of these results that would seem to be a fairly large value. The sensitivities of the Consistency measure scale with the measure, and for low values of the measure the sensitivities are very low.

Theil's θ changed with the confidence of the views and generally indicated that much of the information in the posterior originated with the prior. It moved intuitively with the changes in confidence level.

Across the two scenarios the TEV decreased by 61bps, but against a starting point of 8.28% it still indcates large active weights. It is not clear what a realistic threshold for the TEV is in this case, but these values are likely towards the upper limit that would be tolerated. Note that between the two scenarios, the sensitivity to the first view dropped by 50% which is consistent with the change in confidence that we applied.

In analyzing these various measures of the tilt caused by the views, the TEV of the weights measures the impact of the views and the optimization process, which we can consider as the final outputs. If the investor is concerned about limits on TEV, they could be easily added as constraints on the optimization process.

He and Litterman's Lambda measures the weight of the view on the posterior weights, but only in the case of an unconstrained optimization. This makes it suitable for measuring impact and being a part of the process, but it cannot be used as a constraint in the optimization process.

Theil's compatibility measure and Fusai and Meucci's consistency measure measure the posterior distribution, including the returns and the covariance matrix. The former in view space the latteri in asset space.

Active Management and the Black-Litterman Model

We will consider Active Management to be the case when our investor is managing an overlay portfolio on top of a 100% invested passive benchmark portfolio. In this case we are only interested in the overlay, and not in the benchmark so we start with a prior distribution with 0 active weights and 0 expected excess returns over the benchmark. All returns to views input to the model are relative to the equilibrium benchmark returns rather than to the risk free rate. The weights are for the active portfolio, so the weights should always sum to 0.

Herold (2003) discusses the application of the Black-Litterman model to the problem of Active Management. He introduces a measure Φ which is He and Litterman's Λ_A modified for Active Management. When we use the Black-Litterman model for Active Management versus a passive benchmark portfolio, then the equilibrium weights (w_{eq}) are 0, and the thus the equilibrium returns (Π) are also 0.

Because the value $\Pi = 0$ the middle term in He and Litterman's Λ_A disappears.

(66) $$\Phi = \frac{1}{\delta}\left[\Omega^{-1}Q - (P\Sigma P^T + \Omega)^{-1} P\Sigma P^T \Omega^{-1} Q\right]$$

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 (6), but add an additional term for the new market factor(s).

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

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
δ_0 is the risk aversion parameter of the market
δ_j is the risk aversion parameter for the j-th additional risk factor
β_j is the vector of exposures to the j-th additional risk factor

Given their utility function as shown in formula (67) we can take the first derivative with respect to w in order to solve for the equilibrium asset returns.

$$(68) \qquad \Pi = \delta_0 \Sigma w + \sum_{j=1}^{n} \delta_j \beta_j$$

Comparing this to formula (7), the simple reverse optimization formula, we see that the equilibrium excess return vector (Π) is a linear composition of (7) and a term linear in the β_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 (68) by v and solving for δ_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 (68) by v_j and substitute δ_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 δ_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 (67) 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)

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

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

thus

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

We can directly substitute $\hat{\Pi}$ and $\hat{Q}$ into formula (25) 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.

Qian and Gorman

Qian and Gorman (2001) discuss a method to provide both a conditional mean estimate and a conditional covariance estimate. They use a Bayesian framework and reference the Black-Litterman model. They use the alternative reference model as τ does not appear in their paper and they neglect the conditional (or posterior) covariance estimate.

In this section we will compare the Qian and Gorman approach with the approach taken in the Black-

Litterman Reference Model.

We can match the variance portion of Qian and Gorman's formula (15) our label (QG 15) with our formula (30) if we set Ω = 0, and remove τ (this is the alternative reference model). This describes the scenario where the investor has 100% confidence in their views. For those assets where the investor has absolute views then the variance of the posterior estimate will be zero. For all other assets, the posterior variance will be non-zero.

(QG 15) $$\epsilon \sim N\left[0, \Sigma - (\Sigma P^T)(\Sigma_v)^{-1}(P\Sigma)\right]$$

where $\Sigma_v = P \Sigma P^T$

$$Var(\epsilon) = \Sigma - (\Sigma P^T)(P \Sigma P^T)^{-1}(P \Sigma)$$

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

setting Ω = 0 and removing τ

$$M = \Sigma - \Sigma P^T (P\Sigma P^T)^{-1} P\Sigma$$

In order to get to Qian and Gorman's formula 17 (QG 17), we need to re-introduce the covariance of the views, Ω, but rather than mix the covariances as is done in the Black-Litterman model, we rely on their lack of correlation and just add the two variance terms. We take the variance of the conditional from (13).

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

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

(69) $$Var(\epsilon) = \Sigma + (\Sigma P^T)((P\Sigma P^T)^{-1}(\Omega)(P\Sigma P^T)^{-1} - (P\Sigma P^T)^{-1})(\Sigma P^T)$$

Note that formula (69) exactly matches (QG 17).

(QG 17) $$\tilde{\Sigma} = \Sigma + (\Sigma P^T)(\Sigma_v^{-1}\tilde{\Sigma}_v\Sigma_v^{-1} - \Sigma_v^{-1})(P\Sigma)$$

where $\Sigma_v = P\Sigma P^T$ and $\tilde{\Sigma}_v = \Omega$. substituting

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

Qian and Gorman demonstrate a conditional covariance which is not derived from either Theil's mixed estimation nor Bayesian updating. It's behavior may increase the variance of the posterior vs the prior in the event that the view variance is larger than the prior variance. They describe this as allowing the investor to have a view on the variance, and they do not suggest that Ω needs to be diagonal. In this model the conditional covariance is proportional to the investor's views on variance, but the blending process is not clear.

Compare this formula to our formula (30) for the variance of the posterior mean estimate.

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

Intuition tells us that the updated posterior estimate of the mean will be more precise (lower variance) than the prior estimate, thus we would like the posterior (conditional) variance to always be less than the prior variance. The Black-Litterman posterior variance achieves this goal and arrives at the well known result of Bayesian analysis in the case of unknown mean and known precision. As result, it seems we should favor these results over Qian and Gorman.

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.

Meucci (2006) and Meucci (2008) provide further extensions to the Black-Litterman Model for non-normal views and views on parameters other than return. This allows one to apply the Black-Litterman Model to new areas such as alternative investments or derivatives pricing. His methods are based on simulation and do not provide a closed form solution. Further analysis of his extensions will be provided in a future revision of this document.

An Asset Allocation Process Using the Black-Litterman Model

When used as part of an asset allocation process, the Black-Litterman model provides for estimates which lead to more stable and more diversified portfolios than estimates derived from historical returns when used with unconstrained mean-variance optimization. Because of this property an investor using mean-variance optimization is less likely to require artificial constraints to get a portfolio without extreme weights. Unfortunately using this model requires a broad variety 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. Then, they need to estimate a covariance matrix for the excess returns of the assets. This is most often done using historical data for an appropriate time window. Both Litterman (2003), and Bevan and Winkelmann (1998) provide details on the process used to compute covariance matrices at Goldman Sachs. In the literature, monthly covariance matrices are most commonly estimated from 60 months of historical excess returns. If the actual asset return itself cannot be used, then an appropriate proxy can be used, e.g. S&P 500 Index for US Domestic Large Cap equities. The return on a short term sovereign bond, e.g US 4 or 13-week treasury bill, would suffice for most United States investor's risk free rate.

When applied to the asset allocation problem, 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, commodities, even institutional investors might have a difficult time piecing together adequate market capitalization information. Return data for these same asset classes can also be complicated by delays, smoothing and inconsistencies in reporting. Further complicating the problem is the question of how to deal with hedge funds or absolute return managers. The question of whether they should be considered a separate asset class is beyond the scope of this paper.

Next, the investor needs to quantify their views so that they can be applied and new return estimates computed. The views can be derived from quantitative or qualitative processes, and can be complete or incomplete, or even conflicting.

Finally, the outputs from the model need to be fed into a portfolio selection model 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 of the prior to market factors besides the asset returns. It is fairly simple to extend the Black-Litterman model to use a multi-factor model for the prior distribution. Krishnan and Mains (2005) have provided extensions to the model which allow adding additional cross asset factors which are not priced in the market. Examples of such factors are a recession, or credit, market factor. Their approach is general and could be applied to other factors if desired.

Most of the Black-Litterman literature reports results using the closed form solution for unconstrained mean variance optimization. They also tend to use non-extreme views in their examples. We believe this is done to keep the papers simple, 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. portfolio selection is normally considered as part of a two step process, first compute the optimal portfolio, and then determine position along the Capital Market Line.

For the ensuing discussion, we will refer to 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.

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
- Determine the market capitalization for each asset class.
- 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. Furthermore, separating public real estate holdings from equity holdings (e.g. REITS in the S&P 500 index) may also be required. Idzorek (2006) provides an example of the analysis required to include commodities as an asset class.

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 increased weights to more recent data and less to older data. Other filtering (Random Matrix Theory) or shrinkage methods could also be used in an attempt to impart additional stability to the process.

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 δ 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 δ. For their international fixed income investments they used an expect Sharpe Ratio of 1.0 for the market. The investor then needs to calibrate τ in some manner. This value is usually on the order of 0.025 ~ 0.050.

At this point almost all of the machinery is in place. The investor needs to specify views on the market. These views can impact one or more assets, in any combination. The views can be consistent, or they can conflict. An example of conflicting views would be merging opinions from multiple analysts, where they may not all agree. The investor needs 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 consistent with their reference model and measured by one of the methods discussed previously.

Appendix D shows the process of cranking through formulas (25), (30) and (27) 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.

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.

Beach, Steven and Orlov, Alexei (2006), "An Application of the Black-Litterman Model with EGARCH-M-Derived views for International Portfolio Management", Working paper.

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

Black, Fischer and Litterman, Robert (1990), "Asset Allocation: Combining Investor Views with Market Equilibrium", Goldman Sachs Fixed Income Research Note, September, 1990.

Black, Fischer and Litterman, Robert (1991a), "Global Portfolio Optimization", *Journal of Fixed Income*, Vol 1, No2, 7-18.

Black, Fischer and Litterman, Robert (1991b), "Global Asset Allocation with Equities, Bonds and Currencies", Goldman Sachs Fixed Income Research Note, October, 1991.

Black, Fischer and Litterman, Robert (1992) "Global Portfolio Optimization.", *Financial Analysts Journal*, Sept/Oct 1992.

Braga, Maria Debora and Natale, Franceso Paolo (2007). "TEV Sensitivity to Views in Black-Litterman Model".

Christadoulakis (2002) Bayesian Optimal Portfolio Selection: The Black-Litterman Approach. Class Notes.

Da, Zhi and Jagnannathan, Ravi (2005), Teaching Note on Black-Litterman Model.

DeGroot (1970) Optimal Statistical Decisions, Wiley Interscience.

Firoozye Nick and Blamont, Daniel (2003), "Asset Allocation Model", Global Markets Research, Deutsche Bank, July 2003.

Frost, Peter and Savarino, James (1986). "An Empirical Bayes Approach to Efficient Portfolio Selection", *Journal of Financial and Quantitative Analysis*, Vol 21, No 3, September 1986.

Fusai and Meucci, Attilio (2003) "Assessing Views", *Risk Magazine*, **16**, 3, S18-S21.

Giacometti, Rosella, Bertocchi, Marida, Rachev, Svetlozar T. and Fabozzi, Frank J. (2007) "Stable distributions in the Black-Litterman approach to asset allocation", *Quantitative Finance*, 7:4, 423 - 433

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

Herold, Ulf (2003), “Portfolio Construction with Qualitative Forecasts”, *Journal of Portfolio Management*, Fall 2003, p61-72.

Herold, Ulf (2005), “Computing Implied Returns in a Meaningful Way”, *Journal of Asset Management*, Vol 6, 1, 53-64.

Idzorek, Thomas (2005), “A Step-By-Step guide to the Black-Litterman Model, Incorporating User-Specified Confidence Levels”, Working paper.

Idzorek, Thomas (2006), “Strategic Asset Allocation and Commodities”, Ibbotson White Paper.

Koch (2004) Consistent Asset Return Estimates. “The Black-Litterman Approach”. Cominvest presentation, October 2004.

Krishnan, Hari and Mains, Norman (2005), “The Two-Factor Black-Litterman Model.”, *Risk Magazine*, July 2005.

Litterman, et al (2003). Modern Investment Management: An Equilibrium Approach by Bob Litterman and the Quantitative Research Group, Goldman Sachs Asset Management, Chapter 7.

Mankert, Charlotta (2006), “The Black-Litterman Model – Mathematical and Behavioral Finance Approaches Towards its Use in Practice”, 2006. Licentiate Thesis.

Martellini, Lionel and Ziemann, Volker (2007), “Extending Black-Litterman Analysis Beyond the Mean-Variance Framework”, *Journal of Portfolio Management*, Summer 2007, Vol 33, 4, 33-44.

Meucci, Attilio (2005), Risk and Asset Allocation, 2005. Springer Finance.

Meucci, Attilio (2006), “Beyond Black-Litterman in Practice: A Five-Step Recipe to Input Views on non-Normal Markets”. Working paper available on SSRN.

Meucci, Attilio (2008). “Fully Flexible Views: Theory and Practice”, Working paper available on SSRN.

Meucci, Attilio (2010). “The Black-Litterman Approach: Original Model and Extensions”. Working paper available on SSRN.

Qian, Edward and Gorman, Stephen (2001). “Conditional Distribution in Portfolio Theory”, *Financial Analysts Journal*, September, 2001.

Salomons, Anisa (2007), “The Black Litterman Model, Hype or Improvement?”, Thesis.

Satchell, Stephen and Scowcroft, Alan (2000) “A Demystification of the Black-Litterman Model: Managing Quantitative and Traditional Portfolio Construction”, *Journal of Asset Management*, Vol 1, 2, 138-150..

Theil, Henri (1971), Principles of Econometrics, Wiley.

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 (1971) 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 $\pi = x\beta + u$

Where π is the mean of the prior return distribution, β is the expected return and u is the normally distributed residual with mean 0 and variance Φ.

Next we consider some additional information, the conditional distribution.

A.2 $q = p\beta + v$

Where q is the mean of the conditional distribution and v is the normally distributed residual with mean 0 and variance Ω.

Both Ω and Σ are assumed to be non-singular.

We can combine the prior and conditional information by writing:

A.3 $$\begin{bmatrix}\pi \\ q\end{bmatrix} = \begin{bmatrix}x \\ p\end{bmatrix}\beta + \begin{bmatrix}u \\ v\end{bmatrix}$$

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

$$E\left(\begin{bmatrix}u \\ v\end{bmatrix}\begin{bmatrix}u' & v'\end{bmatrix}\right) = \begin{bmatrix}\Phi & 0 \\ 0 & \Omega\end{bmatrix}$$

We can then apply the generalized least squares procedure, which leads to estimating β as

A.4 $$\hat{\beta} = \left[\begin{bmatrix}x & p\end{bmatrix}\begin{bmatrix}\Phi & 0 \\ 0 & \Omega\end{bmatrix}^{-1}\begin{bmatrix}x' \\ p'\end{bmatrix}\right]^{-1}\begin{bmatrix}x' & p'\end{bmatrix}\begin{bmatrix}\Phi & 0 \\ 0 & \Omega\end{bmatrix}^{-1}\begin{bmatrix}\pi \\ q\end{bmatrix}$$

This can be rewritten without the matrix notation as

A.5 $$\hat{\beta} = \left[x\Phi^{-1}x' + p\Omega^{-1}p'\right]^{-1}\left[x'\Phi^{-1}\pi + p'\Omega^{-1}q\right]$$

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

A.6 $\hat{\beta}=\left[x\Phi^{-1}x'+p\Omega^{-1}p'\right]^{-1}\left[x'\Phi^{-1}(x\beta+u)+p'\Omega^{-1}(p\beta+v)\right]$

This simplifies to

$$\hat{\beta}=\left[x\Phi^{-1}x+p\Omega^{1}p'\right]^{-1}\left[x\beta\Phi^{-1}x+p'\Omega^{-1}p\beta+x\Phi^{-1}u+p\Omega^{-1}v\right]$$

$$\hat{\beta}=\left[x\Phi^{-1}x'+p\Omega^{-1}p'\right]^{-1}\left[x\Phi^{-1}x'\beta+p\Omega^{-1}p'\beta\right]+\left[x\Phi^{-1}x'+p\Omega^{1}p'\right]^{-1}\left[x\Phi^{-1}u+p\Omega^{-1}v\right]$$

$$\hat{\beta}=\beta+\left[x\Phi^{-1}x'+p\Omega^{1}p'\right]^{-1}\left[x\Phi^{-1}u+p\Omega^{-1}v\right]$$

A.7 $\hat{\beta}-\beta=\left[x\Phi^{-1}x'+p\Omega^{1}p'\right]^{-1}\left[x\Phi^{-1}u+p\Omega^{-1}v\right]$

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

$$E((\hat{\beta}-\beta)^2)=(\left[x\Phi^{-1}x^T+p\Omega^{1}p^T\right]^{-1}\left[x\Phi^{-1}u^T+p\Omega^{-1}v^T\right])^2$$

$$E((\hat{\beta}-\beta)^2)=\left[x\Phi^{-1}x^T+p\Omega^{1}p^T\right]^{-2}$$
$$\left[x\Phi^{-1}u^T u\Phi^{-1}x^T+p\Omega^{-1}v^T v\Omega^{-1}p^T+x\Phi^{-1}u^T v\Omega^{-1}p^T+p\Omega^{-1}v^T u\Phi^{-1}x^T\right]$$

We know from our assumptions above that $E(uu')=\Phi$, $E(vv')=\Omega$ and $E(uv')=0$ because u and v are independent variables, so taking the expectations we see the cross terms are 0

$$E((\hat{\beta}-\beta)^2)=\left[x\Phi^{-1}x^T+p\Omega^{-1}p^T\right]^{-2}\left[(x\Phi^{-1}\Phi\Phi^{-1}x^T)+(p\Omega^{-1}\Omega\Omega^{-1^T})+0+0\right]$$

$$E((\hat{\beta}-\beta)^2)=\left[x\Phi^{-1}x^T+p\Omega^{-1}p^T\right]^{-2}\left[x\Phi^{-1}x^T+p\Omega^{-1}p^T\right]$$

And we know that for the Black-Litterman model, *x* is the identity matrix and $\Phi=\tau\Sigma$ so after we make those substitutions we have

A.8 $E((\hat{\beta}-\beta)^2)=\left[(\tau\Sigma)^{-1}+p\Omega^{-1}p^T\right]^{-1}$

Appendix B

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.

A Quick Introduction to Bayes Theory

Bayes theory states

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

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 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

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

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

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

Ω is the uncertainty in the estimate μ of the mean, it is not the variance of the distribution about the mean.

Then the posterior distribution is specified by

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

The variance term in (73) 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 (73) 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 (73).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 (73) in the absence of a conditional distribution, it should collapse into the prior distribution.

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

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

As we can see in formula (74), 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 (73) into a form which is more easy to work with in the 100% certainty case.

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

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

We will revisit equations (73) and (75) later in this paper where we transform these basic equations into the various parts of the Black-Litterman model. Appendix D contains derivations of the alternate Black-Litterman formulas from the standard form, analogous to the transformation from (73) to (75).

Appendix C

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

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

Next, we consider the PDF for the conditional distribution.

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

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

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

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

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

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

C.4 $$y=\frac{\left(\mu\Sigma^{-1}+x(S/n)^{-1}\right)}{\left(\Sigma^{-1}+(S/n)^{-1}\right)}$$

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 proportionality factor and we are left we

C.5 $$\xi(x|\mu)\propto\exp\left(-\left[\left(\Sigma^{-1}+(S/n)^{-1}\right)^{-1}(E(R)-y)^2\right]\right)$$

Thus the posterior mean is y as defined in formula C.4, and the variance is

C.6 $$\left(\Sigma^{-1}+(S/n)^{-1}\right)^{-1}$$

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 (24) we will derive formula (25).

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

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.

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

Appendix E

Derivation of formula (12)

We start with the definition of the views.

(76) $\hat{Q} = Q + \epsilon$

Where

$\hat{Q}$ is the k x 1 vector of the unknown mean returns to the views.
Q is the k x 1 vector of the estimated mean returns to the views.
ϵ is a n x 1 matrix of residuals from the regression where $E(\epsilon)=0 ; V(\epsilon)=E(\epsilon\epsilon^T)=\Omega$ and Ω is non-singular.

We can rewrite formula (1) into a distribution of $\hat{Q}$ as follows:

(77) $\hat{Q} \sim N(Q, \Omega)$

We can also write our definition of the the unknown mean returns of the views based on the unknown mean returns of the assets and the portfolio pick matrix P

(78) $P\hat{\Pi} = \hat{Q}$

Where

P is the k x n vector of weights for the view portfolios.
$\hat{\Pi}$ is the n x 1 vector of the unknown returns of the assets

Substituting (1) into (2) we get the following

(79) $P\hat{\Pi} = Q + \epsilon$

Assuming that P is invertible, which requires it to be of full rank then we can multiply both sides by P^{-1}. This is the projection of the view estimated means into asset space representing the Black-Litterman conditional distribution. If P is not invertible then we would need a slightly different formulation here, adding another term on the right hand side.

(80) $\hat{\Pi} = P^{-1}Q + P^{-1}\epsilon$

We would like to represent formula (5) as a distribution. In order to do this we need to compute the covariance of the random term. The variance of the unknown asset means about the estimated view means projected into asset space is calculated as follows:

(81)

$$
\begin{aligned}
&Variance = E\left[P^{-1}\epsilon\epsilon^{T}\left[P^{-1}\right]^{T}\right]\\
&Variance = P^{-1}E\left[\epsilon\epsilon^{T}\right]\left[P^{-1}\right]^{T}\\
&Variance = P^{-1}\Omega\left[P^{-1}\right]^{T}\\
&\quad using\left[P^{T}\right]^{-1} = \left[P^{-1}\right]^{T}\\
&Variance = P^{-1}\Omega\left[P^{T}\right]^{-1}\\
&\quad using\left[AB\right]^{-1} = B^{-1}A^{-1}\\
&Variance = \left[P^{T}\Omega^{-1}P\right]^{-1}
\end{aligned}
$$

So we arrive at the projection of the views into asset space as

(82) $$\hat{\Pi} \sim N\left(P^{-1}Q, \left[P^{T}\Omega^{-1}P\right]^{-1}\right)$$

The covariance term here is the covariance of the unknown mean returns about the estimated returns from the views, it is not the covariance of expected returns.

Appendix F

This section of the document summarizes the steps required to implement the Black-Litterman model.

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 (7).

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

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 ω_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 (25).

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

We compute the posterior variance using formula (30).

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

Closely followed by the computation of the sample variance from formula (27).

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

And now we can compute the portfolio weights for the optimal portfolio on the unconstrained efficient frontier from formula (10).

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