Mathematical Derivations and Practical Implications for the Use of the Black-Litterman Model

Executive Summary. *In this article, the financial portfolio model often referred to as the Black-Litterman model is described, and then mathematically derived, using a sampling theoretical approach. This approach generates a new interpretation of the model and gives an interpretable formula for the mystical parameter, τ, the weight-on-views. The practical implications of the model are discussed, along with how portfolio fund managers should arrive at model input values and what consideration must be weighted beforehand.*

by Charlotta Mankert*
Michael J. Seiler**

In 1952, Markowitz published the article *Portfolio Selection,* which is the origin of modern portfolio theory. Portfolio models are tools intended to help portfolio managers determine the weights of the assets within a fund or portfolio. Markowitz's ideas have had a great impact on portfolio theory and have withstood the test of time. However, in practical portfolio management the use of Markowitz's model has not had the same impact as it has had in academia. Many fund and portfolio managers consider the composition of the portfolio given by the Markowitz model as unintuitive (Michaud, 1989; Black and Litterman, 1992). The practical problems in using the Markowitz model motivated Black and Litterman (1992) to develop a new model in the early 1990s. The model, often referred to as the Black-Litterman model (hereafter the B-L model), builds on Markowitz's model and aims at handling some of its practical problems. While optimization in the Markowitz model begins from the null portfolio, the optimization in the B-L model begins from, what Black and Litterman refer to as, the equilibrium portfolio (often assessed as the benchmark weights of the assets in the portfolio). "Bets" or deviations from the equilibrium portfolio are then taken on assets to which the investor has assigned views. The manager assigns a level of confidence to each view indicating how sure he/she is of that particular view. The level of confidence affects how much the weight of that particular asset in the B-L portfolio differs from the weights of the equilibrium portfolio.

*KTH Royal Institute of Technology, SE-100 44 Stockholm or charlotta.mankert@indek.kth.se.

**Old Dominion University, Norfolk, VA 23529-0223 or mseiler@odu.edu.

The purpose of this article is to (1) carefully and methodologically describe and mathematically derive the B-L model, (2) review the relevant literature to discuss the model's implications to practical implementation of its usage, (3) present and discuss theoretical starting points for future research, and (4) establish a foundation for the discussion of Mankert and Seiler (2011).

The Markowitz Model

Markowitz (1952) focuses on a portfolio as a whole, instead of an individual security selection when identifying an optimal portfolio. Previously, little research concerning the mathematical relations within portfolios of assets had been carried out. Markowitz began from John Burr Williams' *Theory of Investment Value*. Williams (1938) claimed that the value of a security should be the same as the net present value of future dividends. Since the future dividends of most securities are unknown, Markowitz claimed that the value of a security should be the net present value of *expected* future returns. Markowitz claims that it is not enough to consider the characteristics of individual assets when forming a portfolio of financial securities. Investors should take into account the co-movements represented by covariances of assets. If investors take covariances into consideration when forming portfolios, Markowitz argues they can construct portfolios that generate higher expected return at the same level of risk or a lower level of risk with the same level of expected return than portfolios ignoring the co-movements of asset returns. Risk, in Markowitz' model (as well as in many other quantitative financial models) is assessed as the variance of the portfolio. The variance of a portfolio in turn depends on the variance of the assets in the portfolio and on the covariances between the assets.

A detailed review of Markowitz (1952) is provided in Appendix 1. A summary of the model is provided in this section, with focus on the practical problems encountered in its use. The practical problems in using the model prompted Black and Litterman (1990, 1992) to continue the development of portfolio modeling.

Markowitz (1952) shows that investors under certain assumptions can, *theoretically*, build portfolios that maximize expected return given a specified level of risk, or minimize the risk given a level of expected return. The model is primarily a normative model. The objective for Markowitz has not been to explain how people select portfolios, but how they should select portfolios (Sharpe, 1967). Even before 1952, diversification was a well-accepted strategy to lower the risk of a portfolio, without lowering the expected return, but until then, no thorough foundation existed to validate diversification. Markowitz's model has remained to date the cornerstone of modern portfolio theory (Elton and Gruber, 1997).

The Model

According to Markowitz (1952), inputs needed to create optimal portfolios include expected returns[1] for every asset, variances for all assets, and covariances between all of the assets handled by the model. In his model, investors are assumed to want as high an expected return as possible, but at as low a risk as possible. This seems quite reasonable. There may be many other factors that investors would like to consider, but this model focuses on risk and return.

The following problem needs to be solved in order to derive the set of attainable portfolios (derived from the expected return and the covariance matrix estimated by the investor) that an investor can reach:

$$\begin{cases} \min_{\mathbf{w}} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \\ \mathbf{w}^T \bar{\mathbf{r}} = \bar{r}_p \end{cases} \tag{1}$$

or

$$\begin{cases} \max_{\mathbf{w}} \mathbf{w}^T \bar{\mathbf{r}} \\ \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} = \sigma_p^2 \end{cases} \tag{2}$$

where:

$\mathbf{w}$ = The column vector of portfolio weights;
$\mathbf{w}^*$ = The Markowitz' optimal portfolio;

σ_p^2 = The variance of the portfolio;
$\bar{r}_p$ = The expected return of the portfolio;
$\bar{\mathbf{r}}$ = The column vector of expected returns;
$\boldsymbol{\mu}$ = The column vector of expected (excess) returns; and
Σ = The covariance matrix.
δ = The risk aversion parameter stated by the investors. States the trade-off between risk and return. Equals μ_P/σ_P^2.[2] This is consistent with Satchell and Scowcroft (2000, p. 139). Economists would call this parameter the standard price of variance.

Often the following problem is solved instead of those above:

$$\max_{\mathbf{w}} \mathbf{w}^T\boldsymbol{\mu} - \frac{\delta}{2}\mathbf{w}^T\Sigma\mathbf{w}. \tag{3}$$

This is actually the same as solving problem (1) or (2) (proof in Appendix 1). Solving these equations generates:

$$\mathbf{w}^* = (\delta\Sigma)^{-1}\boldsymbol{\mu}. \tag{4}$$

This is the formula for Markowitz's (1952) optimal portfolio.

Problems in the Use of the Markowitz Model

Although the model might seem appealing and reasonable from a theoretical point of view, several problems arise when using the model. Michaud (1989) thoroughly discusses the practical problems of using the model. He claims that the model often leads to irrelevant optimal portfolios and that some studies have shown that even equal weighting can be superior to Markowitz optimal portfolios. Michaud argues that the most important reason for many financial actors not to use the Markowitz model is "political." According to Michaud, quantitatively-oriented specialists would have a central role in the investment process and would intimidate more qualitatively-oriented managers, as well as top level managers; however, the article was written 20 years ago and this may no longer be the most important reason for not using the model. Michaud suggests additional disadvantages of using the model.

The most important problems in using the Markowitz model are:

- According to Michaud (1989) and Black and Litterman (1992), the Markowitz model optimizers maximize errors. Since there are no correct and exact estimates of either expected returns or variances and covariances, these estimates are subject to estimation errors. Markowitz optimizers overweight securities with high expected return and negative correlation and underweight those with low expected returns and positive correlation. These securities are, according to Michaud, those that are most prone to large estimation errors. However, the argument appears somewhat contradictory. The reason for investors to estimate a high expected return on assets should be that they believe this asset is prone to high returns. It then seems reasonable that the manager would appreciate that the model overweighs this asset in the portfolio (taking covariances into consideration).
- Michaud (1989) claims that the habit of using historical data to produce a sample mean and replace the expected return with the sample mean is not a good one. He claims that this line of action contributes greatly to the error-maximization of the Markowitz model.
- The Markowitz model does not account for the market capitalization weights of assets. This means that if assets with a low level of capitalization have high expected returns and are negatively correlated with other assets in the portfolio, the model can suggest a high portfolio weight. This is actually a problem, especially when adding a shorting constraint. The model then often suggests very high weights in assets with low levels of capitalization.
- The Markowitz mean-variance model does not differentiate between different levels of

uncertainty associated with the estimated inputs to the model.

- Mean-variance models are often unstable, meaning that small changes in inputs might dramatically change the portfolio. The model is especially unstable in relation to the expected return input. One small change in expected return on one asset might generate a radically different portfolio. According to Michaud (1989), this mainly depends on an ill-conditioned covariance matrix. He exemplifies ill-conditioned covariance matrixes by those estimated with "insufficient historical data."

Michaud (1989) also discusses further problems with the Markowitz model. These are non-uniqueness, exact versus approximate mean-variance optimizers, inadequate approximation power, and default settings of parameters.

One of the most striking empirical problems in using the Markowitz model is that when running the optimizer without constraints, the model almost always recommends portfolios with large negative weights in several assets (Black and Litterman, 1992). Fund or portfolio managers using the model are often not permitted to take short positions. Thus, a shorting constraint is often added to the optimization process. What happens then is that when optimizing a portfolio with constraints, the model gives a solution with zero weights in many of the assets and therefore takes large positions in only a few of the assets and unreasonably large weights in other assets. Many investors find portfolios of this kind unreasonable and although it seems as though many investors are attracted to the idea of mean-variance optimization, these problems appear to be among the main reasons for not using it. In a world in which investors are quite sure about the inputs to an optimization model, the output of the model would not seem so unreasonable. In reality however, every approximation about future return and risk is quite uncertain and the chance that it is "absolutely correct" is low. Since the estimation of future risk and return is uncertain, it seems reasonable that investors wish to invest in portfolios that are not prospective disasters if the estimations prove incorrect. The Markowitz model has been shown, however, to generate portfolios that are very unstable (i.e., sensitive to changes in inputs; Fisher and Statman, 1997), meaning that a small change in input radically changes the structure of the portfolio. Michaud (1989) claims that better input estimates could help bridge problems of the lack of intuitiveness of the Markowitz portfolios. Fisher and Statman, however, maintain that although good estimates are better than bad, better estimates will not bridge the gap between mean-variance optimized portfolios and "intuitive" portfolios, in which investors are willing to invest, since estimation errors can never be eliminated. It is not possible to predict expected returns, variances, and covariances with 100% confidence.

Estimating covariances between assets is also problematic. In a portfolio containing 50 assets, the number of variances that need to be estimated is 50, but the number of covariances that need to be estimated is 1,225. This seems a bit much for a single portfolio manager to handle. It also seems much for an investment team, consisting of several persons. According to Markowitz (1991, p. 102) "in portfolios involving large numbers of correlated securities, variances shrink in importance compared to covariances."

Although there are several severe disadvantages in the use of the Markowitz model, such as the idea of maximizing expected return, minimizing risk or optimizing the trade-off between risk and expected return is so appealing that the search for better-behaved models has continued. The B-L model is one of these, and has gained much interest in recent years.

Historical Data

There seems to exist a common misconception that Markowitz's theories and model build solely on historical data. This is not the case. Markowitz asserts that various types of information can be used as an input to a portfolio analysis.

> "One source of information is the past performance of individual securities. A second source of information is the beliefs of one or more security analysts concerning future performances," (Markowitz, 1991, p. 3).

> "Portfolio selection should be based on reasonable beliefs about future returns rather than past performances per se. Choices based on past performances alone assume, in effect, that average returns of the past are good estimates of the 'likely' return in the future; and variability of return in the past is a good measure of the uncertainty of return in the future," (Markowitz 1991, p. 14).

Markowitz (1991) focuses on portfolio analysis, not security analysis. He claims that he does not discuss how to arrive at a reasonable belief about securities since this is the job of a security analyst. His contribution begins where the contribution of the security analysis ends. While he repeats that historical data alone are inadequate as a basis for estimating future returns and covariances, the importance of historical data in modern financial theory had been well made in the literature. Indeed, it is hard to question the fact that historical time series have had great impact on financial decision making.

> "...covariance matrices determined from empirical financial time series appear to contain such a high amount of noise that their structure can essentially be regarded as random. This seems, however, to be in contradiction with the fundamental role played by covariance matrices in finance, which constitute the pillars of modern investment theory and have also gained industry-wide applications in risk management," (Pafka and Kondor, 2002, Abstract).

There seems to be a general confusion between the covariances of future returns and the covariances estimated from historical data. This is problematic and may affect the discussion and the development of portfolio theory. The discussion whether historical data are a good approximation for future covariance matrices is important. Also, it is important to discuss whether or not it is possible to make reasonable estimates of future covariances and how this affects the use of portfolio modeling. Separating the two discussions would be productive.

The Black-Litterman Model

The problems encountered when using the Markowitz model in practical portfolio management and the fact that mean-variance optimization has not had such a high impact in practice motivated Fisher Black and Robert Litterman to work on the development of models for portfolio choice. Black and Litterman (1992) proposed a means of estimating expected returns to achieve better-behaved portfolio models. However, they require the portfolio to be at the efficient frontier. If this is not the case, it may be possible to obtain a "better" portfolio from a mean-variance perspective. The B-L model is often referred to as a completely new portfolio model. Actually, the B-L model differs only from the Markowitz model with respect to the expected returns. The B-L model is otherwise theoretically quite similar to the Markowitz model. How the B-L expected returns are to be estimated has been found to be quite complicated. The model generates portfolios differing considerably from portfolios generated using the Markowitz model.

The Framework and the Idea

The B-L model was developed to make portfolio modeling more useful in practical investment situations (Litterman, 2003a). Black and Litterman (1992) apply what they call an equilibrium approach to do this. They set the idealized market equilibrium as a point of reference. The investor then specifies a chosen number of market views in the form of expected returns and a level of confidence for each view. The views are combined with the equilibrium returns and the combination of these constitutes the B-L expected returns. The B-L expected returns are then optimized in a mean-variance way, creating a portfolio where bets are taken on assets where investors have opinions about expected returns, but not elsewhere. The size of the bets, in relation to the equilibrium portfolio weights, depends on the confidence levels specified by the user and also on a parameter specifying the weight of the collected investor views in relation to the market equilibrium, the weight-on-views.

The following notation is used:

$\mathbf{w}^*$ = The weight vector of the B-L unconstrained optimal portfolio.
$\mathbf{w}^M$ = The weight vector of the market capitalized portfolio, referred to as the equilibrium portfolio or the market portfolio.
δ = The risk aversion factor. It is, according to Black and Litterman (1990), proportionality constant based on the formulas in Black (1989).
$\delta = \mu_P/\sigma_P^2$ (Satchell and Scowcroft, 2000, p. 139). In He and Litterman (1999), the authors use "$\delta = 2.5$ as the risk aversion parameter representing the world average risk tolerance."
Σ = The covariance matrix containing variances of and covariances between all the assets handled by the model.
$\mathbf{P}$ = A matrix representing a part of the views. Each row in the matrix contains the weights of assets of one view. The maximum number of rows (i.e., the maximum number of views) is the number of assets in the portfolio.
$\overline{\mathbf{q}}$ = A column vector that represents the estimated expected returns in each view.
ω_i = The level of confidence assigned to view i. It is the standard deviation around the expected return of the view so that the investor is 2/3 sure that the return will lie within the interval.
$\boldsymbol{\Omega}$ = A diagonal matrix consisting of $\omega_1^2,\ldots,\omega_k^2$.
τ = A parameter often referred to as the *weight-on-views*. τ is a constant, which together with $\boldsymbol{\Omega}$ determines the weighting between the view portfolio and the equilibrium portfolio.
$\boldsymbol{\mu}^*$ = This is the B-L modified vector of estimated expected returns.
$\boldsymbol{\Pi}$ = The column vector of equilibrium expected excess returns.

To derive the B-L expected returns estimated by the market, the following problem is solved:

$$\max_{\boldsymbol{\Pi}}(\mathbf{w}^M)^T\boldsymbol{\Pi} - \frac{\delta}{2}(\mathbf{w}^M)^T\Sigma\mathbf{w}^M.$$

Equilibrium excess returns, $\boldsymbol{\Pi}$ is:

$$\boldsymbol{\Pi} = \delta\Sigma\mathbf{w}^M. \tag{5}$$

This formula represents the expected returns estimated by the market. Many managers, however, do not wish to invest in the market portfolio. They have views that differ from the market returns. The market returns are then combined with investor views and a modified vector of expected returns constituting the B-L vector of expected returns is created. This new vector of B-L expected returns is then optimized in a mean-variance manner, yielding the formula for the weights of the optimal portfolio. The formula for the Black-Litterman optimal portfolio, without constraints, is presented below. Readers need not understand this formula at this point—a detailed derivation and explanation will be given later in this section. For now, consider the following formula:

$$\mathbf{w}^* = \mathbf{w}^M + \frac{\tau}{\delta}\mathbf{P}^T(\boldsymbol{\Omega} + \tau\mathbf{P}\Sigma\mathbf{P}^T)^{-1}(\mathbf{q} - \delta\mathbf{P}\mathbf{w}^M). \tag{6}$$

The formula implies that the model takes the market weights and then adds a component. Hence the model starts of from the market weights.

Equilibrium

What do Black and Litterman mean by equilibrium? Litterman and the Quantitative Research Group (2003) discuss the concept of the equilibrium approach. Equilibrium, they suggest, is an idealized state in which supply equals demand. This state never actually occurs in financial markets, but they argue that there are a number of attractive characteristics about the idea. There are "natural forces," in the form of arbitrageurs, in the economic system that function to eliminate deviations from equilibrium. Even if there are disturbances in markets, such as noise traders, uncertain information, and lack of liquidity that result in situations in which deviations are large and in which adjustment takes time, there is a tendency that mispricing will, over time, be "corrected." Hence, the markets are not assumed to be in equilibrium (Litterman, 2003b). Equilibrium is instead viewed as a "centre of gravity." Markets deviate from this state, but forces in the system will push markets towards equilibrium. The idea of an equilibrium as a point of reference for the B-L model is a kind of ideal condition for the model. In order to apply the model to real life investment situations, making a reasonable approximation of this state is needed.

Litterman (2003b) claims that the reason for recommending the equilibrium approach is the belief that it is a favorable and appropriate point of reference from which identification of deviations can be made and taken advantage of. He admits that no financial theory can ever capture the complexity of financial markets. Still, "Financial theory has the most to say about markets that are behaving in a somewhat rational manner. If we start by assuming that markets are simply irrational, then we have little more to say," (Litterman, 2003b). He refers to the extensive literature that can be accessed if we are willing to accept the assumption of arbitrage-free markets. According to Litterman, there is also a need to add the assumption that markets, over time, move toward a rational equilibrium in order to take advantage of portfolio theory. He states that portfolio theory makes predictions about how markets will behave, tells investors how to structure their portfolios, how to minimize risk and also how to take maximum advantage of deviations from equilibrium.

Much literature concerning the B-L model assumes a global asset allocation model, and because of this Litterman (2003a) argues that the global Capital Asset Pricing Model (CAPM) is a good starting point for a global equilibrium model. Black (1989) discusses an equilibrium model providing a framework from which the B-L global asset allocation model has emerged. However, the B-L model is not used only in global asset management, but also in domestic equity portfolio management and fixed income portfolio management. In such cases, the equilibrium weights are easier to find by using the domestic CAPM.

There is an obvious problem in using equilibrium weights as a point of reference since these weights are not observable and therefore must be estimated. Bevan and Winklemann (1998) present a way of dealing with this. If the market is in equilibrium, a representative investor will hold a part of the capitalization-weighted portfolio. Many investors are evaluated according to a benchmark portfolio. Often the benchmark is a capitalization-weighted index (Litterman, 2003c). The equilibrium portfolio is then approximated as the benchmark portfolio. These estimated expected returns could be seen as the expected returns estimated by the market if all actors in the market act in a mean-variance manner. Expected equilibrium returns are calculated from the benchmark weights using equation 4. As Schachter, Hood, Andreassen, and Gerin (1986, p. 254) state: "[T]he price of a stock is more than an objective, rationally determined number; it is an opinion, an aggregate opinion, the moment-to-moment resultant of the evaluation of the community of investors." For each asset, to which the investor has no view, this is what will be handed over to the optimizer. For the assets to which the investor has views, modified expected returns are calculated as a combination of the benchmark weights and the investor views. This way of estimating the equilibrium portfolio is what will be used in this section. Henceforth, the equilibrium portfolio will often be referred to as the market portfolio.

Investor Views and Levels of Confidence

The B-L idea is to combine the equilibrium with investor-specific views. For each view, a level of confidence is to be set by the manager. The model allows the investor to express both absolute and relative views. An example of an absolute view is "I expect that equities in country A will return X%." An example of a relative view is "I believe domestic bonds will outperform domestic equities by Y%." In traditional mean-variance portfolio optimization, relative views cannot be expressed. The investor must also assign a level of confidence to each view, whether stated in the relative or absolute form. The level of confidence is expressed as the standard deviation around the expected return of the view. If managers feel confident in one view, the standard deviation should be small and if they are not confident in a view, the standard deviation should be large. The confidence level affects the influence of a particular view. The weaker confidence that is set to a view, the less the view is to affect the portfolio weights. This is considered as an attractive feature since views most often are incorrect. Views, however, indicate on which assets investors want to take bets and in which direction the bets ought to be taken.

Combining Views with the Equilibrium Expected Returns

The B-L optimal portfolio is a weighted combination of the market portfolio and the views of the investor. The views are combined with the equilibrium, and positions are taken in relation to the benchmark portfolio on assets to which investors have expressed views. The size of the bet taken depends on three different variables: the views, the level of confidence assigned to each view, and the weight-on-views. It depends on the views specified by the investor. Views that differs much from the market expected returns contribute to larger bets. If the level of confidence assigned to a view is strong, this also contributes to larger bets. The more confidence the investor assigns to a view, the larger the bets are on that particular asset.

The matrix $\boldsymbol{\Omega}$ represents the levels of confidence of the views. There is, however, one more variable that affects the size of the bets taken in relation to the equilibrium portfolio. The variable τ, the weight-on views (Bevan and Winkelmann, 1998), determines, with $\boldsymbol{\Omega}$, how much weight is to be set on the set of view portfolios specified by the investor in relation to the equilibrium portfolio. There appears to be no clear description of this variable in the existing literature. Black and Litterman (1992, p. 17) propose that the constant should be set close to zero "because the uncertainty in the mean is much smaller than the uncertainty in the return itself." However, Satchell and Scowcroft (2000) claim that τ often is set to 1, but they also claim that this is not always successful in reality. Alternatively, Bevan and Winkelmann (1998), on the other hand, suggest that τ can be set so that the information ratio[3] does not exceed 2.0. They have found that τ most often lies between 0.5 and 0.7. He and Litterman (1999), on the other hand, claim that τ need not be set at all, since only $\tau^{-1}\boldsymbol{\Omega}$ enters the model. Mathematically, this is correct, but then there would be no point in specifying these two variables from the beginning. The reasoning concerning τ is thus quite weak in the literature. The articles do not express any associations to normative and descriptive argumentation. There are totally different suggestions on what τ ought to be set to and explanations of why these are reasonable values of τ is not properly given. Later an interpretable formula to the weight-on-views will be derived and explained. One of the great advantages of taking a sampling theoretical approach to the B-L model is that it provides an interpretable formula to the weight-on-views. While we do not state a definitive recommended value for τ, the formula will give the user of the B-L model guidance in setting this variable.

When no investor views are specified, the B-L model recommends holding the market portfolio. If investors have no opinion about the market they should not place bets in relation to the equilibrium weights. However, if they have opinions about assets, it seems reasonable that the bets are placed in those assets and the rest of the assets have weights close to the market-capitalized portfolio. The stronger confidence assigned to both the individual view and the weight-on-views, the more the output portfolio deviates from the market portfolio.

The Bayesian Approach to the B-L Model

Most of the literature concerning the B-L model makes use of a Bayesian[4] approach to construe the model. The approach combines prior information (information considered as relevant although not necessarily in the form of sample data) with sample data. Through repeated use of Bayes' theorem,[5] the prior information is updated. Although the Bayesian approach to inference is conceptually quite different from the sampling theory approach to inference, the results of the two methods are nearly identical. An example of an important difference between the approaches is that in the sampling theory approach, we consider θ, the estimate of the unknown parameter μ, to be an unknown constant, while the Bayesian approach views θ as a random variable.

As mentioned, the most frequent way of interpreting the B-L model is from a Bayesian point of view. Since the idea is to update information from the market with information from the investor, the Bayesian approach lays easy at hand. Two papers that clearly use the Bayesian approach are Satchell and Scowcroft (2000) and Christodoulakis and Cass (2002). Satchell and Scowcroft (2000, p. 139) claim that the B-L model is, in fact, based on a Bayesian methodology and also that this "methodology effectively updates currently held opinions with data to form new opinions." The authors point out that despite the importance of the model, it appears, as if there is no comprehensible description of the mathematics underlying the model.

In the Bayesian approach, we need to decide what is to be considered as prior information and what is to be considered as sample information. Satchell and Scowcroft (2000) use investor views as prior information and information from the market is seen as sample data with which they update the investor views to receive the posterior distribution. Satchell and Scowcroft admit that their interpretation of what is prior information and what is the sample data *may* differ from that of others. It might be questioned whether this is a good way to *demystify* the B-L model. The authors also claim that the aim of Black and Litterman was to form a model that made the idea of combining investor views with market equilibrium sensible to investors. We argue that neither Black and Litterman nor Satchel and Scowcroft have succeeded with this task. If Black and Litterman had produced a text that made the idea of combining investor views with the market equilibrium comprehensible to investors, there would be no need for Satchell and Scowcroft to write an article intended to demystify the model. Satchell and Scowcroft, however, assert that the Bayesian approach has been undermined by the problems in specifying a numerical distribution representing the view of an individual. It is claimed in the article that the parameter τ is a "known scaling factor that often is set to one," (p. 140). The parameter is not explained in any further way.

Christodoulakis and Cass (2002) also interpret the B-L model in a Bayesian manner. They claim that the articles by Black and Litterman provide more of a framework for combining investor views with the market equilibrium, than a sensible and clear description of the model. The authors argue, consistent with Satchell and Scowcroft (2000), for using investor views as the prior information and market equilibrium returns for updating these to receive the posterior expected returns. The fact that the model assumes investor views are formed independently of each other is discussed. This is implied by the assumption that the returns are normally distributed together with the fact that $\boldsymbol{\Omega}$ is a diagonal matrix. The B-L model assumes a diagonal $\boldsymbol{\Omega}$-matrix. However, this is an inconsistency in the model. Christodoulakis and Cass (2002, p. 5) refer to τ as a scalar known to the investor that scales the "historical covariance matrix Σ." That they refer to Σ as the historical covariance matrix is questionable. Our interpretation of the B-L model is that Σ is the same covariance matrix as that in the Markowitz model and neither Markowitz nor Black and Litterman claim that this should be anything else than the estimated future covariances between the assets that the model handles.

Sampling Theory Approach and the Black-Litterman Model

So far, the explanation of the B-L model has focused on the idea and the framework behind it. Some parts of this model are difficult to understand on the basis of existing literature. Papers by He and Litterman (1999), and Satchell and Scowcroft (2000), and Idzorek (2004) suggest that others have encountered such problems with the model as well. It seems relatively easy to grasp the framework, but understanding how the formula for the B-L vector of expected returns is derived is quite a challenge. The papers addressing the B-L model begin from a Bayesian perspective. The idea of trying to derive the model from a sampling theory point of view was actually presented by Black and Litterman (1992, pp. 34–35):

> "One way we think about representing that information is to act as if we had a summary statistic from a sample of data drawn from the distribution of future returns—data in which all that we're able to observe is the difference between the returns of A and B. Alternatively, we can express this view directly as a probability distribution for the difference between the means of the excess returns of A and B. It doesn't matter which of these approaches we want to use to think about our views; in the end we get the same result."

While most people seem to have chosen the Bayesian approach, the quotation implies that the authors also had the sampling theoretical approach in mind.

Sampling Theory[6]

In sampling theory, the study of sample data is supposed to shed light on an unknown parameter. The unknown parameter can, for example, be the variance or the expected value of a stochastic variable or stochastic vector. Point estimation is a well-known concept of sampling theory. Point estimates of the expected excess returns are what we want to estimate in the B-L model.

Different realizations of the stochastic variable or the vector of stochastic variables may generate different values yielding different estimates. The resulting probability distribution is called the sampling distribution of the statistic. Sampling theory cannot yield statements of final precision. Let us clarify this with an example. Consider dice of uncertain symmetry (i.e., with an unknown probability function). The true probability function of the dice is $p_Y(y)$. The sampling theoretical way of getting information of the probability function of the dice is by throwing them a number of times and studying the results. Here the sample data, y, is represented by the outcome of one toss of the dice. Different tosses will generate different outcomes $y_1, y_2, \ldots, y_n$. If the unknown parameter, μ, is the expected value, we can estimate this by calculating its sample mean. The sample mean could then act as an estimate θ of the unknown parameter μ. The sample mean is:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

All estimates of unknown parameters are not accurate. According to sampling theory, characteristics that an estimator should possess include freedom from bias, consistency, sufficiency, efficiency, and low variance, etc.[7] One of the most recognized methods for point estimation within sampling theory is the maximum likelihood method. The estimates generated by this method possess many of the characteristics of a good estimator.

"Likelihood," or rather the likelihood function, is a central element within classical statistics. It is simply a re-interpretation of the density function and expresses how the probability (density), $p_\theta(x)$, of the data x fluctuates with different values of the parameter θ. However, the likelihood is not the probability (density) of θ for a given sample x. The likelihood function for θ based on the sample data $X = x$ is given by:

$$L_\theta(x) = p_\theta(x_1) \cdot p_\theta(x_2) \cdot \ldots \cdot p_\theta(x_n).$$

The maximum likelihood method is a statistical method for estimating parameters from sample data. The parameter values that maximize the probability of obtaining the observed data are selected as estimates. The method is one of the most widely used for constructing estimators. The estimates resulting from this approach often possess the desirable properties originating from the classical approach. Maximum likelihood methods possess many attractive features (Barnett, 1999).

In general, the idea is that the value of the parameter under which the obtained data would have had highest probability (density) of occurring must be the best estimator of θ. Intuitively, we can think of the estimates as the value of θ that best supports the observed sample.

When working with the maximum likelihood method we almost always choose to work with the logarithm of the likelihood function instead of the likelihood function itself. We do this because the log-likelihood function is easier to work with and both the likelihood function and the log-likelihood function have their maximum values for the same θ. We differentiate the log-likelihood function, set it equal to zero, and solve for θ to obtain the maximum likelihood estimator. The maximum likelihood method generates good estimators when we have a good model for the underlying distributions and their dependence of the parameter θ. A poor model for underlying distributions may, not surprisingly, generate bad estimates.

Although the classical approach to inference seems to make sense and is widely applied, it has its critics. Criticism is focused on two fundamental factors within sampling theory. The first is the preoccupation with a frequency-based probability concept providing justification for assessing the behavior of statistical procedures in terms of their long-term behavior. The criticism questions the validity of assigning aggregate properties to specific inferences. The second type of criticism of sampling theory relates to the restrictions applied by the approach on what is regarded as relevant information, namely sample data (Barnett, 1999).

The Sampling Theory Approach to the Black-Litterman Model

One reason for trying a sampling theoretical approach to the B-L model has to do with the problems experienced when trying to gain a deeper understanding of the model from the existing literature. Since sampling theory is just another way of considering inference and point estimation, the idea of using the approach appears interesting. Sampling theory builds on sample data as information for inference, but in this case, we have no sample data. The two approaches, Bayesian and sampling theory, will however be seen to generate the same result.

We suppose that both the market and the individual investor have observed samples of future returns to handle the fact that we have no sample data. The sample returns observed by the market will then represent the equilibrium portfolio, while the sample returns observed by the investor will represent the investor's views. The samples observed by the market are different from those observed by the investor.

Suppose that the market has observed a number of samples of future asset returns. Using the method of maximum likelihood, we derive the markets' estimated expected returns, referred to as the equilibrium or market returns. We also suppose that the investor has observed a number of samples of returns. The investor has observed returns on a number of portfolios of assets instead of on the assets themselves. These portfolios can relate to all the assets in the investor universe or just one or a few of them. We use the maximum likelihood method to estimate the expected returns of the investor's views. We assume that the observations of future asset returns are normally distributed.

The Equilibrium Portfolio

Suppose the market has observed m samples of asset returns and that the investment universe contains d assets. Then suppose the market has observations in the following form:

$$\mathbf{r}_1 = \begin{bmatrix} r_1^1 \\ \vdots \\ r_1^d \end{bmatrix}, \mathbf{r}_2 = \begin{bmatrix} r_2^1 \\ \vdots \\ r_2^d \end{bmatrix}, \ldots, \mathbf{r_m} = \begin{bmatrix} r_m^1 \\ \vdots \\ r_m^d \end{bmatrix}.$$

From these, we derive the market estimated expected returns and equilibrium returns:

$$\Pi = \bar{\mathbf{r}}^M = \begin{bmatrix} \bar{r}^1 \\ \vdots \\ \bar{r}^d \end{bmatrix} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{r}_i,$$

by using the method of maximum likelihood. Assume that the observed samples of the market are drawn from a normal distribution with the true vector of expected value equal to $\boldsymbol{\mu}$ and the covariance matrix equal to Σ. Then the vector of sample means is normally distributed with the vector of expected returns, $\boldsymbol{\mu}$, and the covariance matrix, Σ/m, i.e.:

$$\mathbf{r}_i \in N(\boldsymbol{\mu}, \Sigma), \; i = 1 \ldots m.$$

$$\bar{\mathbf{r}}^M \in N\left(\boldsymbol{\mu}, \frac{\Sigma}{m}\right).$$

The probability function of the return is then:

$$p(\mathbf{r}_i) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{r}_i - \boldsymbol{\mu})\right).$$

Since we are only interested in for which value of $\boldsymbol{\mu}$ the likelihood function (i.e., the product of the probability functions) takes its maximum value, we do not need to consider the constants. Instead we will work with:

$$\varphi(\mathbf{r}_i) = \exp\left(-\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{r}_i - \boldsymbol{\mu})\right).$$

The likelihood function is then:

$$\mathbf{L} = \varphi(\mathbf{r}_1) \cdot \varphi(\mathbf{r}_2) \cdot \ldots \cdot \varphi(\mathbf{r}_m).$$

As previously mentioned, the logarithm of the likelihood function is easier to work with and the log-likelihood function is then:

$$\ell = \ln\mathbf{L} = \ln[\varphi(\mathbf{r}_1)\cdot\varphi(\mathbf{r}_2)\cdot\ldots\cdot\varphi(\mathbf{r}_m)] = \ln\varphi(\mathbf{r}_1) + \ln\varphi(\mathbf{r}_2) + \ldots + \ln\varphi(\mathbf{r}_m).$$

$$\left\{\ln\varphi(\mathbf{r}_i) = \ln\left[\exp\left(-\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu})\right)\right] = -\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{r}_i - \boldsymbol{\mu})\right\}.$$

$$\ell = \frac{1}{2}\left(-\sum_{i=1}^{m}(\mathbf{r}_i - \boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu})\right).$$

We want to maximize the log-likelihood function:

$$\max_{\boldsymbol{\mu}} \ell = \max_{\boldsymbol{\mu}} \frac{1}{2}\left(-\sum_{i=1}^{m}(\mathbf{r}_i - \boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu})\right)$$

Let us differentiate the function with respect to $\boldsymbol{\mu}_j$ and set the derivative equal to zero. We use the notation:

$$\mathbf{e}_j = [0\ldots\underset{\uparrow \text{ entry } j}{1}\ldots0],\ m \text{ elements}$$

$$\frac{\partial}{\partial\boldsymbol{\mu}^j}\ell = \frac{1}{2}\sum_{i=1}^{m}(-\mathbf{e}_j^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu}^{*M}) - (\mathbf{r}_i - \boldsymbol{\mu}^{*M})^T\Sigma^{-1}\mathbf{e}_j) = 0$$

$$\{(\mathbf{r}_i - \boldsymbol{\mu}^{*M})^T\Sigma^{-1}\mathbf{e}_j = [(\mathbf{r}_i - \boldsymbol{\mu}^{*M})^T\Sigma^{-1}\mathbf{e}_j]^T = \mathbf{e}_j^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu}^{*M})\}$$

$$-\mathbf{e}_j^T\Sigma^{-1}\sum_{i=1}^{m}(\mathbf{r}_i - \boldsymbol{\mu}^{*M}) = 0$$

$$-\mathbf{e}_j^T\Sigma^{-1}\left(\sum_{i=1}^{m}\mathbf{r}_i - \sum_{i=1}^{m}\boldsymbol{\mu}^{*M}\right) = 0$$

$$m\mathbf{e}_j^T\Sigma^{-1}(\overline{\mathbf{r}}^M - \boldsymbol{\mu}^{*M}) = 0$$

Since this holds for all $j = 1,\ldots,d$, it follows that:

$$\boldsymbol{\mu}^{*M} = \overline{\mathbf{r}}^M = \frac{1}{m}\sum_{i=1}^{m}\mathbf{r}_i.$$

$$\boldsymbol{\Pi} = \boldsymbol{\mu}^{*M}.$$

$\boldsymbol{\mu}^{*M}$ is hence the expected future excess return estimated by the market.

The Views of the Manager

Let us assume that an investor has observed n other samples of returns. These observations are not observations of returns on individual assets, but rather the returns on portfolios of assets. As described above, the investor need not state views about every asset in his investment universe. Instead, a number of portfolios are chosen and the investor postulates that he observes a number of samples of the future returns of these portfolios. The weights of the portfolios are expressed in a matrix, **P**, in which each position represents the weight of a certain asset in a certain view portfolio. Each row in the matrix represents one view portfolio and for each view portfolio the investor expresses an expected return $\overline{q}_i$ and a level of confidence ω_i. Suppose the investor has opinions about k portfolios, $k \leq d$, where d is the number of assets handled by the model. In the B-L model, **P** is the matrix:

$$\mathbf{P} = \begin{bmatrix} w_1^1 & \ldots & w_1^d \\ \vdots & \ddots & \vdots \\ w_k^1 & \ldots & w_k^d \end{bmatrix},$$

where w_j^i is the weight of asset i in view portfolio j.

The expected returns to each portfolio are referred to as:

$$\overline{\mathbf{q}} = \begin{bmatrix} \overline{q}_1 \\ \vdots \\ \overline{q}_k \end{bmatrix}$$

where $\overline{\mathbf{q}} = \mathbf{P}\overline{\mathbf{r}}^I$.

From this formula, we can derive the expected returns to each asset estimated by the investor:

$$\overline{\mathbf{r}}^I = \mathbf{P}^{-1}\overline{\mathbf{q}}.$$

To clarify how to set **P** and $\overline{\mathbf{q}}$, consider an example of the two easiest and perhaps most used views. Consider a portfolio holding just three assets: A, B, and C. As such, the investor can express three or fewer views. In this example, only two views are expressed: View 1: I believe asset A will return 3%;

View 2: I believe asset B will outperform asset C by 2%. $\mathbf{P}$ and $\bar{\mathbf{q}}$ will then appear as follows:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \bar{\mathbf{q}} = \begin{bmatrix} 3\% \\ 2\% \\ 0 \end{bmatrix}$$

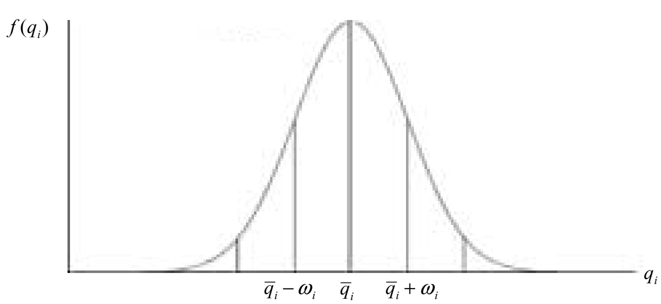
Each row in $\mathbf{P}$ represents one view portfolio. Each column represents the weights of a specific asset.

The diagonal matrix represents the investor's levels of confidence $\boldsymbol{\Omega}$. $\omega_1^2,\ldots,\omega_k^2$ constitute the diagonal of $\boldsymbol{\Omega}$. The number of rows and columns equals the number of views stated by the investor.

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_1^2 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \omega_{\kappa}^2 \end{bmatrix}$$

The possibility to express a level of confidence to each view is, to many, considered to be the most attractive feature of the B-L model. But what is a level of confidence? How is this supposed to be estimated? Let us remind ourselves of the samples of portfolio returns observed by the investor. We assumed that the investor had observed n samples of the returns of the view portfolios and that the samples were normally distributed. The level of confidence, ω_i^2, is the variance of $\bar{q}_i$. ω_i can be interpreted as an interval around $\bar{q}_i$, so that two-thirds, of the postulated samples lie within the interval $\bar{q}_i \pm \omega_i$, where $i = 1,\ldots,k$, see Exhibit 1.

Exhibit 1
Level of Confidence



Note: The level of confidence, ω_i^2, is the variance of $\bar{q}_i$. ω_i can be interpreted as an interval around $\bar{q}_i$, so that two-thirds of the postulated samples lie within the interval $\bar{q}_i \pm \omega_i$, where $i = 1,\ldots,k$.

The samples observed by the investor are also supposed to be drawn from a normally distributed set. The vector of expected values is the same as for the market (i.e., $\boldsymbol{\mu}$). The covariance matrix, however, is not the same.

$$\underbrace{\mathbf{r}_1,\ldots,\mathbf{r}_m,}_{\substack{m \text{ observation} \\ \text{by the market}}} \quad \underbrace{\mathbf{r}_{m+1},\ldots,\mathbf{r}_{m+n}}_{\substack{n \text{ observations} \\ \text{by the investor}}}$$

Since $\mathbf{r}_j \in N(\boldsymbol{\mu}, \Sigma)$ and $\mathbf{q}_j = \mathbf{P}\mathbf{r}_j$, then $\mathbf{q}_j$ should be $N(\mathbf{P}\boldsymbol{\mu}, \mathbf{P}^T\Sigma\mathbf{P})$.[8] However, in the B-L model, the distribution of $\mathbf{q}_j$ is $\mathbf{q}_j \in N(\mathbf{P}\boldsymbol{\mu}, \boldsymbol{\Omega})$. Hence, this is an inconsistency since $\boldsymbol{\Omega} \neq \mathbf{P}^T\Sigma\mathbf{P}$. $\boldsymbol{\Omega}$ is a diagonal matrix implying that returns on the portfolios observed by the investor are uncorrelated. This is an inconsistent assumption because the returns on the assets from which the portfolios are formed has the covariance matrix Σ, and Σ is not diagonal.

We do not derive the maximum likelihood estimator of the investor observations. The procedure is the same as for the market with the only difference being the number of observations. The market has observed m samples and the investor has observed n samples. The maximum likelihood estimator of the expected excess return of the investor is:

$$\boldsymbol{\mu}^{*I} = \bar{\mathbf{q}} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{q}_j = \frac{1}{n}\sum_{j=1}^{n} \mathbf{P}\bar{\mathbf{r}}_j$$
$$= \mathbf{P}\frac{1}{n}\sum_{j=1}^{n} \mathbf{r}_j = \mathbf{P}\bar{\mathbf{r}}^I.$$

Combining Investor Views with Market Equilibrium

We now derive the maximum likelihood estimator of the expected returns from the returns observed by the market together with the returns observed by the investor.

$$\max_{\boldsymbol{\mu}} \sum_{i=1}^{m} -\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu})$$
$$+ \sum_{j=m+1}^{m+n} -\frac{1}{2}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu})^T\boldsymbol{\Omega}^{-1}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}).$$

We will use:

$$\mathbf{e}_k = [0\ldots010\ldots0], \; n + m \text{ elements.}$$

entry k

Let us differentiate the function with respect to $\boldsymbol{\mu}_j$ and set the derivative equal to zero.

$$\frac{\partial}{\partial \boldsymbol{\mu}^k}\left(\sum_{i=1}^{m} -\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu}^*)^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu}^*) + \sum_{j=m+1}^{m+n} -\frac{1}{2}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}^*)^T\boldsymbol{\Omega}^{-1}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}^*)\right)$$

$$\frac{1}{2}\sum_{i=1}^{m}(-\mathbf{e}_k^T\Sigma^{-1}(\mathbf{r}_i - \boldsymbol{\mu}^*)^T - (\mathbf{r}_i - \boldsymbol{\mu}^*)\Sigma^{-1}\mathbf{e}_k) + \frac{1}{2}\sum_{j=m+1}^{m+n}(-\mathbf{e}_k^T\mathbf{P}\boldsymbol{\Omega}^{-1}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}^*)^T - (\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}^*)\boldsymbol{\Omega}^{-1}\mathbf{P}\mathbf{e}_k) = 0.$$

$$e_k^T\Sigma^{-1}\sum_{i=1}^{m}(\mathbf{r}_i - \boldsymbol{\mu}^*) + e_k^T\mathbf{P}\boldsymbol{\Omega}^{-1}\sum_{j=m+1}^{m+n}(\mathbf{q}_j - \mathbf{P}\boldsymbol{\mu}^*) = 0.$$

$$e_k^T(m\Sigma^{-1}(\boldsymbol{\Pi} - \boldsymbol{\mu}^*) + n\mathbf{P}\boldsymbol{\Omega}^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\mu}^*)) = 0.$$

Since this is true for all $k = 1,\ldots,n + m$, we get:

$$\frac{m}{n}\Sigma^{-1}(\boldsymbol{\Pi} - \boldsymbol{\mu}^*) + \mathbf{P}\boldsymbol{\Omega}^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\mu}^*) = 0.$$

We then set:

$$\tau = \frac{n}{m}.$$

$$\boldsymbol{\mu}^*(\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P} + \tau^{-1}\Sigma^{-1}) = \mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}} + \tau^{-1}\Sigma^{-1}\boldsymbol{\Pi}.$$

$$\boldsymbol{\mu}^* = [(\tau\Sigma)^{-1} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1} + [(\tau\Sigma)^{-1}\boldsymbol{\Pi} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}}].$$

This gives us the B-L formula for the modified vector of expected returns:

$$\boldsymbol{\mu}^* = [(\tau\Sigma)^{-1} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}\cdot[(\tau\Sigma)^{-1}\boldsymbol{\Pi} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}}]. \quad (7)$$

This is the form most often used in the literature. Another way of expressing the B-L vector of modified expected returns is:[9]

$$\boldsymbol{\mu}^* = \boldsymbol{\Pi} + \tau\Sigma\mathbf{P}^T(\boldsymbol{\Omega} + \tau\mathbf{P}\Sigma\mathbf{P}^T)^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}). \quad (8)$$

This way of presenting the B-L modified vector of expected returns may appear as more intuitive than the original formula. We see here that the modified vector of expected returns consists of first the vector of expected returns estimated by the market, $\boldsymbol{\Pi}$, and then another expression $\tau\Sigma\mathbf{P}^T(\boldsymbol{\Omega} + \tau\mathbf{P}\Sigma\mathbf{P}^T)^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi})$. Thus, the expected returns estimated by the market are updated with another expression. If the last part of (8), $(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi})$ equals zero (i.e., if the view of the investor is the same as the market view), then the modified vector of the expected return is only $\boldsymbol{\Pi}$. It is not obvious, however, that equation (7) and equation (8) are equal, and it is not at all easy to deduce expression (8) of the modified vector of expected returns from expression (7). We therefore will show how this is done.

$$\begin{aligned}
\boldsymbol{\mu}^* &= [(\tau\Sigma)^{-1} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}\cdot[(\tau\Sigma)^{-1}\boldsymbol{\Pi} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}}] \\
&= [(\tau\Sigma)^{-1} + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}(\tau\Sigma)^{-1}(\tau\Sigma)[(\tau\Sigma)^{-1}\boldsymbol{\Pi} \\
&\quad + \mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}}] \\
&= [\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}\cdot[\boldsymbol{\Pi} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\overline{\mathbf{q}}] \\
&= [\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}\cdot[(\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})\boldsymbol{\Pi} \\
&\quad + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi})] \\
&= \boldsymbol{\Pi} + (\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}\cdot(\tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1} \\
&\quad (\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi})) \\
&= \boldsymbol{\Pi} + (\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}\cdot\tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1} \\
&\quad \{(\boldsymbol{\Omega} + \mathbf{P}^T\tau\Sigma\mathbf{P})(\boldsymbol{\Omega} + \mathbf{P}^T\tau\Sigma\mathbf{P})^{-1}\}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}) \\
&= \boldsymbol{\Pi} + (\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1} \\
&\quad (\tau\Sigma\mathbf{P}^T + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P}^T\tau\Sigma\mathbf{P}) \\
&\quad (\boldsymbol{\Omega} + \mathbf{P}^T\tau\Sigma\mathbf{P})^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}) \\
&= \boldsymbol{\Pi} + (\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}(\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})\tau\Sigma\mathbf{P}^T \\
&\quad (\boldsymbol{\Omega} + \mathbf{P}^T\tau\Sigma\mathbf{P})^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}) \\
&= \boldsymbol{\Pi} + (\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})^{-1}(\mathbf{I} + \tau\Sigma\mathbf{P}^T\boldsymbol{\Omega}^{-1}\mathbf{P})\tau\Sigma\mathbf{P}^T \\
&\quad (\boldsymbol{\Omega} + \mathbf{P}^T\tau\Sigma\mathbf{P})^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}).
\end{aligned}$$

Here one parenthesis is multiplied by its own inverse. Hence we get:

$$\boldsymbol{\mu}^* = \boldsymbol{\Pi} + \tau\Sigma\mathbf{P}^T(\boldsymbol{\Omega} + \tau\mathbf{P}\Sigma\mathbf{P}^T)^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi})$$

or

$$\boldsymbol{\mu}^* = \boldsymbol{\Pi} + \Sigma\mathbf{P}^T\left(\frac{\boldsymbol{\Omega}}{\tau} + \mathbf{P}\Sigma\mathbf{P}^T\right)^{-1}(\overline{\mathbf{q}} - \mathbf{P}\boldsymbol{\Pi}).$$

Using the formula:

$$\mathbf{W}^* = (\delta\Sigma)^{-1}\boldsymbol{\mu}^*$$

we get:

$$\mathbf{W}^* = \mathbf{W}^M + \mathbf{P}^T\left(\frac{\boldsymbol{\Omega}}{\tau} + \mathbf{P}\Sigma\mathbf{P}^T\right)^{-1}\left(\frac{\overline{\mathbf{q}}}{\delta} - \mathbf{P}\Sigma\mathbf{W}^M\right),$$

representing the unconstrained optimal portfolio.

The derivation of the B-L model from the sampling theoretical approach is hereby completed. We have arrived at the same formula for the B-L modified expected returns as reached in papers taking a Bayesian approach. The formula for the B-L modified expected returns are also reformulated and the formula for the weights of the optimal unconstrained portfolio is shown as well.

Readers may wonder whether this approach is really new. Have these calculations not been published previously? Black and Litterman already suggested this method in 1992. However, after an extensive Web search it appears that the sampling theoretical derivations of the B-L model have not been published before.

Results

A Detailed Derivation of the B-L Model from a Sampling Theoretical Approach

It has been shown that the sampling theory approach offers an alternative way to derive the B-L model. The derivation leads to the same formula for the B-L modified vector of expected return as obtained by using a Bayesian approach.

A New Way to Interpret the Model. The sampling theory approach provides a new way to interpret the B-L model. Sampling theory depends solely on sample data, but since we have no sample data, users are required to postulate a number of sample returns. Investors postulate that the market has observed a number of samples of asset returns and that they themselves have observed a number of samples of returns on portfolios of assets. The number of observations need not be specified, but the number of samples observed by the investor in relation to the number of samples observed by the market must be estimated.

A Formula for the Parameter τ, the Weight-on-Views. The derivation has generated a formula for τ as:

$$\tau = \frac{n}{m},$$

where n represents the number of samples observed by the investor and m represents the number of samples observed by the market. Hence, τ is the ratio between these numbers, and it is only this ratio that needs to be estimated. If investors postulate the number of samples they have observed to be the same as the number of samples observed by the market, then τ should equal 1. If investors postulate the numbers of samples observed by the investor to be more numerous than the number of samples observed by the market, τ should be larger than one and vice versa. So, the more confident investors are in all the views, the higher τ should be set.

As previously discussed, it appears that there is no clear description of the variable τ in the literature. Hopefully the sampling theory approach presented here will help investors set τ and help academics as well as practitioners continue the process of testing and further developing the B-L model.

A New Interpretation of the Matrix $\boldsymbol{\Omega}$. The sampling theory approach to the B-L model generates an interpretation of the matrix $\boldsymbol{\Omega}$ that differs somewhat from the Bayesian approach. The level of confidence in an expected return on view i is seen as the value of ω_i^2 so that one standard deviation, about two-thirds of the postulated observed samples of a certain view portfolio lie within the interval $q_i \pm \omega_i$. Note also that investors need not postulate how many samples they have observed; they need only postulate a confidence interval around the expected return of the portfolio so that two-thirds of the postulated samples lie within this interval. However, it is possible to implement the model so that investors estimate both an interval and another percentage. The investor could then, for instance, claim that he believes that in 90% of the n trials, the true return on the view will lie within the interval $q_i \pm \gamma_i . \omega_i^2$ that is then calculated from these data.

An Inconsistency in the Distribution of $\mathbf{q}_j$. The distribution of $\mathbf{q}_j$ is $\mathbf{q}_j \in N(\mathbf{P}\boldsymbol{\mu}, \boldsymbol{\Omega})$. But for the model to be consistent, the distribution should be $N(\mathbf{P}\boldsymbol{\mu}, \mathbf{P}^T\Sigma\mathbf{P})$. Those trying to understand the B-L model should benefit from knowing of this inconsistency. If unaware it is probable that people will be confused, believing that there is something they have misunderstood. It will probably be easier to handle $\boldsymbol{\Omega}$ knowing of this inconsistency.

Conclusion

It is our hope that the results might contribute to a more thorough understanding of the B-L model. New ways of deriving models should constitute a contribution both to academics and practitioners. A derivation of the B-L model from a sampling theoretical approach hopefully facilitates understanding of the B-L model by individuals not familiar or comfortable with Bayesian theory. The fact that the approach generates an interpretable formula for τ, the weight-on-views, should also contribute to the development of the model. How would it be possible to understand, use, test, and/or evaluate a model consisting of one parameter of which no clear and interpretable description exists? However, the practical contribution of this derivation will not be known until it is tested "in use." Studying the use of the B-L model can generate knowledge about how users relate to this interpretation.

Appendix 1

Markowitz' Mean-Variance Model

According to Markowitz (1952), the inputs needed to create optimal portfolios are expected returns[10] for every asset, variances for all assets, and covariances between all of the assets handled by the model. Markowitz does not state exactly how these parameters should be estimated, although his discussion of some alternatives is quite detailed. He sees past performance as one source of information, but emphasizes that portfolio selection solely based on historical data assumes that past data are a reasonable approximation of the future. Instead, he prefers the "probability beliefs" of experts as inputs to the portfolio analysis (Markowitz, 1991). He compares the way a security analyst arrives at probability beliefs with the way a meteorologist arrives at a weather forecast and calls the security analyst the meteorologist of stocks and bonds. But, he also emphasizes that portfolio analysis begins where security analyses ends. In Markowitz's model, expected returns are to be estimated as the expected return of every asset during the investment period. Investors specify the length of the investment period.

Risk, in the Markowitz model, as well as in many other financial models, is approximated by the variances and covariances of future returns. When considering only one asset, it is sufficient to estimate and evaluate only its expected future return and the future variance. When evaluating a portfolio of assets, however, we should consider how the assets within the portfolio co-vary to be able to estimate the variance of the portfolio as a whole. The covariance is a measure of how the values of two random variables move up and down together. In this case, the random variables are any pair of assets in a portfolio. The covariance is crucial to portfolio theory and increases the possibilities of getting a well-diversified portfolio.

Choosing a Portfolio

In portfolio theory, investors are assumed to want as high expected return as possible, but at a risk as low as possible. There are many other factors, which investors might consider, but risk and return are what this model focuses on.

We use the following notation:

$\mathbf{w}$ = The column vector of portfolio weights;
$\mathbf{w}^*$ = The Markowitz' optimal portfolio;
σ^2 = The variance of the portfolio;
$\bar{r}_i$ = The expected return of asset number i;

r_{rf} = The return of the risk-free asset;
$\bar{r}$ = The expected return of the portfolio;
w_{rf} = The weight of the risk-free asset as a percentage of the portfolio as a whole;
$\boldsymbol{\mu}$ = The column vector of expected (excess) returns;
Σ = The covariance matrix; and
δ = The risk aversion parameter stated by the investors. It states the trade-off between risk and return.

We set:

$$\bar{\mathbf{r}} = \begin{bmatrix} \bar{r}_1 \\ \vdots \\ \bar{r}_d \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

hence we get:

$$\mathbf{e}r_{rf} = \begin{bmatrix} r_{rf} \\ \vdots \\ r_{rf} \end{bmatrix}.$$

We need to solve the following problem to derive the set of attainable portfolios (derived from the expected return and the covariance matrix estimated by the investor) that an investor can reach:

$$\begin{cases} \min\limits_{\mathbf{w}} \mathbf{w}^T\Sigma\mathbf{w} \\ \mathbf{w}^T\bar{\mathbf{r}} = \bar{r} \end{cases} \tag{A.1}$$

or

$$\begin{cases} \max\limits_{\mathbf{w}} \mathbf{w}^T\bar{\mathbf{r}} \\ \mathbf{w}^T\Sigma\mathbf{w} = \sigma^2 \end{cases} \tag{A.2}$$

We minimize the variance of the portfolio given a certain level of expected return or we maximize the expected return of the portfolio for a certain level of risk (variance).

Exhibit 2 shows the function of all attainable portfolios (i.e., combinations of expected return and, in this case, standard deviation). All combinations to the right of the curve are attainable, whereas those at the left of the curve are not. The combinations on the curve are called the minimum variance set, since for every level of expected return, the point on the curve represents the minimum variance attainable. The upper part of the curve is called the efficient set or the efficient frontier. Portfolios on this part of the curve are referred to as efficient since they represent portfolios generating maximum expected return for a given level of risk. For every portfolio positioned on the lower half of the curve, it is possible to choose other portfolios that stochastically dominate through higher expected return and/or lower risk. Because of this, all portfolios not lying on the efficient frontier are called inefficient.

Exhibit 2
The Efficient Frontier, Attainable Set, and Minimum Variance Portfolios

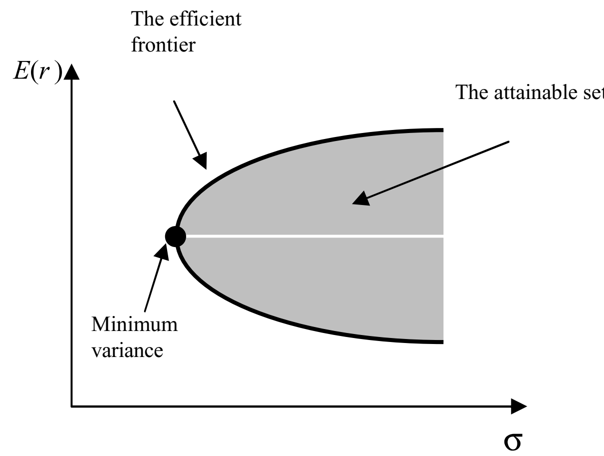
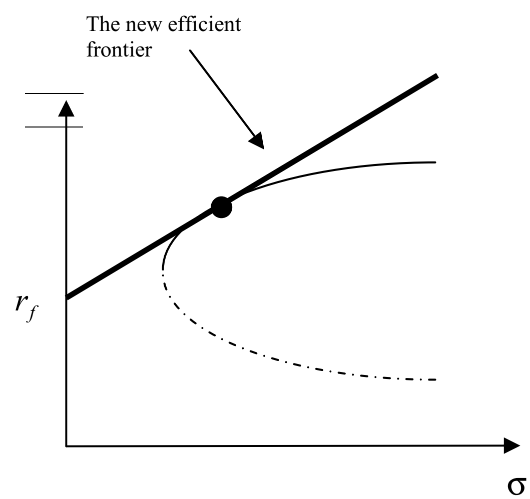


Exhibit 3
Security Market Line



Let us now include a risk-free asset.[11] Assume we have d risky assets. The weight of the risk-less asset in the portfolio is hence:

$$w_{r_f} = 1 - \mathbf{e}^T\mathbf{w}.$$

The expected return of the portfolio, r_P is then:

$$\bar{r}_P = \mathbf{w}^T\bar{\mathbf{r}} + w_{r_f} r_{r_f},$$

and we can write the expected return as:

$$\bar{r}_P = \mathbf{w}^\mathbf{T}\bar{\mathbf{r}} + (1 - \mathbf{w}^\mathbf{T}\mathbf{e})r_{r_f} = \mathbf{w}^\mathbf{T}(\bar{\mathbf{r}} - \mathbf{e}r_{r_f}) + r_{r_f}.$$

We define the vector of expected (excess) returns as:

$$\boldsymbol{\mu} \equiv \bar{\mathbf{r}} - \mathbf{e}r_{r_f} = \begin{bmatrix} \bar{r}_1 - r_{r_f} \\ \vdots \\ \bar{r}_d - r_{r_f} \end{bmatrix}.$$

Hence now the universe of available portfolios has been expanded, and the efficient frontier is moved.

The new efficient frontier is a weighted combination of the risk-free asset and the portfolio in which a straight line drawn from the risk-free rate or return is a tangent to the efficient frontier when no risk-free asset is available. This is also quite reasonable because, in this model, we always want an expected return as high as possible when taking a certain level of risk or as low level of risk as possible for a certain level of expected return.

Let us introduce the parameter δ, often referred to as the risk-aversion parameter. This parameter is a measure of the risk-return trade-off. We are to solve the following problem:

$$\max_{\mathbf{w}} r_{r_f} + \mathbf{w}^T\boldsymbol{\mu} - \frac{\delta}{2}\mathbf{w}^T\Sigma\mathbf{w}.$$

Since r_{r_f} is constant, we can exclude it and still get the same result. The problem to be solved is hence:

$$\max_{\mathbf{w}} \mathbf{w}^T\boldsymbol{\mu} - \frac{\delta}{2}\mathbf{w}^T\Sigma\mathbf{w}. \qquad \text{(A.3)}$$

This problem is solved by setting:

$$\mathbf{e}_k^T = [00...010...0], \text{ number of elements equals number of assets.}$$
$$\uparrow_{\text{entry } k}$$

Differentiate the function and set it equal to zero:

$$\mathbf{e}_k^T\boldsymbol{\mu} - \frac{\delta}{2}\mathbf{e}_k^T\Sigma\mathbf{w} - \frac{\delta}{2}\mathbf{w}^T\Sigma\mathbf{e}_k = 0.$$

$$\mathbf{e}_k^T(\boldsymbol{\mu} - \delta\Sigma\mathbf{w}) = 0.$$

This is true for all $k = 1,...,d \Rightarrow$

$$\mathbf{w}^* = (\delta\Sigma)^{-1}\boldsymbol{\mu}, \qquad \text{(A.4)}$$

where $\mathbf{w}^*$ represents the Markowitz optimal portfolio given the risk aversion coefficient, covariance matrix, and vector of expected returns estimated by the investor.

Problem (A.4) is actually the same as solving problem (A.1). Hence:

$$\begin{cases} \max_{\mathbf{w}} \mathbf{w}^T\boldsymbol{\mu} \\ \mathbf{w}^T\Sigma\mathbf{w} = \sigma^2 \end{cases}$$

The Lagrange function is then:

$$L = \mathbf{w}^T\boldsymbol{\mu} - \lambda(\mathbf{w}^T\Sigma\mathbf{w} - \sigma^2).$$

Differentiate and we get:

$$\mathbf{e}_k^T\boldsymbol{\mu} - 2\lambda\mathbf{e}_k^T\Sigma\mathbf{w} = 0.$$

This is the same as differentiating (A.3), which is:

Let:

$$\lambda = \frac{\delta}{2}.$$

Then:

$$\mathbf{e}_k^T\boldsymbol{\mu} - \delta\mathbf{e}_k^T\Sigma\mathbf{w} = 0.$$
$$\boldsymbol{\mu} = \delta\Sigma\mathbf{w}.$$
$$\mathbf{w} = (\delta\Sigma)^{-1}\boldsymbol{\mu}.$$
$$\sigma^2 = \mathbf{w}^T\Sigma\mathbf{w} = \delta^{-2}\boldsymbol{\mu}^T\Sigma^{-1}\Sigma\Sigma^{-1}\boldsymbol{\mu} = \delta^{-2}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}.$$

This shows that when we select the value of the parameter σ, the value of δ is given. We can also choose a value of δ and we then get the value of σ.

$$\lambda = \frac{\delta}{2}.$$

$\delta/2$ is thus just the Lagrange multiplier.

When:

$$\begin{aligned}
\boldsymbol{\mu} &= \delta\Sigma\mathbf{w} \\
\mathbf{w}^* &= (\delta\Sigma)^{-1}\boldsymbol{\mu} \\
\mu_P &= \mathbf{w}^{*T}\boldsymbol{\mu} = \boldsymbol{\mu}^T(\delta\Sigma)^{-1}\boldsymbol{\mu} = \delta^{-1}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} \\
\sigma_P^2 &= \mathbf{w}^{*T}\Sigma\mathbf{w}^* = \boldsymbol{\mu}^T(\delta\Sigma)^{-1}\Sigma(\delta\Sigma)^{-1}\boldsymbol{\mu} \\
&= \delta^{-2}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} = \delta^{-1}\delta^{-1}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} = \delta^{-1}\mu_P
\end{aligned}$$

then:

$$\delta = \frac{\mu_P}{\sigma_P^2}.$$

This is also consistent with Satchell and Scowcroft (2000). Economists would call this parameter the standard price of variance.

Hence the Markowitz optimized portfolio is:

$$\mathbf{w}^* = (\delta\Sigma)^{-1}\boldsymbol{\mu}.$$

Appendix 2
B-L Model Assumptions

It is not easy to compile a complete list of the B-L model assumptions since many of the assumptions are implicit. As such, the list presented below is not meant to be exhaustive. Instead, it presents key assumptions to keep in mind when using the B-L model.

Assumptions common to many quantitative *financial* models:

- Returns are normally distributed.
- Investors are rational.
- Absence of arbitrage.
- Decreased marginal utility of wealth.
- Increased risk is considered as negative.
- Increased expected return is considered as positive.
- There is a trade-off between expected return and risk.
- Capital markets are efficient in that the prices of securities reflect all available information and further that prices of individual securities adjust very rapidly to new information;

Assumptions common for quantitative *portfolio* models:

- Each possible investment has a probability distribution of expected returns over some holding period.
- Only risk and expected return are used in investment decisions.
- Investors will choose the combination of asset weights that generates the highest expected return for a given risk level. Or, investors will choose the combination of asset weights that generates the lowest risk for a given level of expected return.
- The investor is risk averse.
- A portfolio's risk can be measured by the future variance of and the covariance between the assets' rate of return.
- Taxes and other transaction costs are not taken into account.

Assumptions specific to the B-L model:

- Investors have views about assets that they believe can lead to a better portfolio.
- The market is not completely efficient (Litterman, 2003).
- Investment positions should only be taken in an asset that investors have views.
- Funds or portfolios are evaluated according to a benchmark portfolio.
- For every opinion, a level of confidence can be estimated.

- Investors are not absolutely certain on any view.

Endnotes

1. For simplicity, expected return will refer to the expected excess return over the one-period risk-free rate.
2. For a derivation, see Appendix 1.
3. A risk measure, measuring how well a fund is paid for the active risk taken, hence how much extra the fund returns by deviating from the index portfolio.
4. The theory of Bayesian inference rests primarily on Bayes' theorem. Thomas Bayes' contribution to the literature on probability theory was only two papers published in *Philosophical Transactions* in 1763–1764. Still, his work has had a major impact on probability theory and the theory of statistics. Both papers where published after his death, and there is still some disagreement on exactly what he was suggesting in the second article, entitled "Essay." There are, however, aspects within the articles that are widely agreed upon: the use of continuous frameworks rather than discrete, the idea of inference (essentially estimation) through assessing the chances that an informed guess about a practical situation will be correct, and in proposing a formal description of what is meant by prior ignorance.
5. $P(A/B) = P(B|A)/P(B)\ P(A)$.
   The prior information that is to be entered into a Bayesian model is represented by a probability $P(A)$, the prior probability. This information is then updated by the information of B, which is supposed to be sample data, and represented in the form of likelihood. The resulting probability is referred to as the posterior probability. However, there are two well-known difficulties within the Bayesian theory of inference. First, there is a problem in the interpretation of the probability idea in a particular Bayesian analysis. Second, it is often difficult to specify a numerical representation of the prior probabilities used in the analysis. How do we proceed when the quantities $P(A|B)$ and $P(B|A)$ are unknown? In a Bayesian framework, we would answer that the best we can do is to compute the quantities with all the information we have at our disposal. The central problem in Bayesian theory is how to use a sample drawn independently according to the fixed, but unknown, probability distribution $P(B)$ to determine $P(A|B)$.
6. Sampling theory is the classical approach to statistical inference. It relies solely on sample data, which is represented by their likelihood (Garthwaite, Thomas, Dawson, and Stoddart, 2002). Sampling theory is considered as the classical approach to inference since historically it was the first of three approaches to take form. The other approaches to inference often discussed are the Bayesian approach, which is the common approach to apply to the B-L model. The third approach to inference is the Decision Theory approach. The Decision Theory approach to inference will not be discussed here. Development of sampling theory can be traced to the early 1800s, while the rivalling approaches have taken form during the last 50 years.
7. See Barnett (1999) for additional information about the characteristics of estimators.
8. Some articles actually suggest $\mathbf{q}_j \in N(\mathbf{P}\boldsymbol{\mu}, \mathbf{P}^T\boldsymbol{\Sigma}\mathbf{P})$. This is mathematically correct, however it impairs one of the main ideas of the B-L model, namely that the investor can specify the confidence in each view portfolio.
9. This was brought to our attention by F. Armerin.
10. For simplicity, expected return will refer to the expected excess return over the one-period risk-free rate.
11. Because there is no real risk-free asset, often a short-term government Treasury security is used as a proxy.

References

Barnett, V. *Comparative Statistical Inference*. Third Edition. New York: Wiley, 1999.

Bevan, A. and K. Winkelmann. Using the Black-Litterman Global Asset Allocation Model: Three Years of Practical Experience. *Fixed Income Research*, Goldman Sachs & Co, 1998. www.hss.caltech.edu/media/from-carthage/filer/263.pdf.

Black, F. Universal Hedging: How to Optimize Currency Risk and Reward in International Equity Portfolios. *Financial Analyst Journal*, 1989, July/August, 16–22.

Black, F. and R. Litterman. Asset Allocation: Combining Investors Views with Market Equilibrium. *Fixed Income Research*. Goldman Sachs, 1990.

——. Global Portfolio Optimization. *Financial Analysts Journal*, 1992, September–October, 28–43.

Christodoulakis, G.A. and J.C. Cass. Bayesian Optimal Portfolio Selection: the B-L Approach. *Notes for Quantitative Asset Pricing MSc Mathematical Trading and Finance*. 2002. http://scholar.google.com/url?sa=U&q=http://www.staff.city.ac.uk/~gchrist/Teaching/QAP/optimalportfoliobl.pdf.

Elton, E.J. and M.J. Gruber. Modern Portfolio Theory, 1950 to Date. *Journal of Banking & Finance*, 1997, 21, 1743–59.

Fisher, K.L. and M. Statman. The Mean-Variance-Optimization Puzzle: Security Portfolios and Food Portfolios. *Financial Analysts Journal*, 1997, 53:4, 41–50.

Garthwaite, D.G., M.R. Thomas, A. Dawson, and H. Stoddart. Arable Crops in Great Britain 2002. *Pesticide Usage Survey Report*, 2002, 187, DEFRA, UK

He, G. and R. Litterman. The Intuition Behind Black-Litterman Model Portfolios. Technical report, Goldman Sachs Investment Management Series, Fixed Income Research. December 1999. Available at SSRN: http://ssrn.com/abstract=334304.

Idzorek, T. A Step-by-Step Guide to the Black-Litterman Model. Working paper, 2004. http://www.economia.unimore.it/pattarin_francesco/didattica/papers/idzorek_2004.pdf.

Litterman, R. Beyond Equilibrium, the Black-Litterman Approach. In Litterman et al. (2003) *Modern Investment Management—An Equilibrium Approach*. New Jersey: Wiley, 2003a, 76–88.

——. Introduction: Why an Equilibrium Approach? In Litterman et al. (2003) "*Modern Investment Management—An Equilibrium Approach*. New Jersey: Wiley, 2003b, 3–8.

——. Global Equilibrium Expected Returns. In Litterman et al. (2003) *Modern Investment Management—An Equilibrium Approach.* New Jersey: Wiley, 2003c, 76–88.

Litterman, R. and the Quantitative Research Group. *Modern Investment Management—An Equilibrium Approach*. New Jersey: Wiley, 2003.

Mankert, C. and M.J. Seiler. Behavioral Finance and its Implication in the use of the Black-Litterman Model. *Journal of Real Estate Portfolio Management*, 2011, forthcoming.

Markowitz, H. Portfolio Selection. *Journal of Finance*, 1952, 7:1, 77–91.

——. *Portfolio Selection.* Oxford: Blackwell, 1991.

Michaud, R. The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analyst Journal*, 1989, 45:1, 31–42.

Pafke, S. and I. Kondor. Noisy Covariance Matrices and Portfolio Optimization II. *Physica A*, 2002, 319, 487–94. http://arXiv.org/abs/cond-mat/0205119.

Satchell, S. and A. Scowcroft. A Demystification of the B-L model: Managing Quantitative and Traditional Portfolio Construction. *Journal of Asset Management*, 2000, 1:2, 138–50.

Schachter, S., D.C. Hood, P.B. Andreassen, and W. Gerin. Aggregate Variables in Psychology and Economics: Dependence and the Stock Market. In B. Gilad and S. Kaish (eds.), *Handbook of Behavioral Economics*. Vol. B. Greenwich, CT: JAI Press, 1986, 237–72.

Sharpe, W.F. Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, 1967, 2:2, 76–84.

Williams, J.B. *Theory of Investment Value.* Amsterdam: North-Holland Publishing Company, 1938.