Black-Litterman and Beyond: The Bayesian Paradigm in Investment Management

Petter N. Kolm, Gordon Ritter and Joseph Simonian

January 19, 2021

Petter N. Kolm is a clinical professor and director of the Mathematics in Finance Master's Program at the Courant Institute of Mathematical Sciences, New York University, NY, USA.
petter.kolm@nyu.edu

Gordon Ritter is an adjunct professor at the Courant Institute of Mathematical Sciences, NYU Tandon School of Engineering, Baruch College, Columbia University and General Partner/CIO at Ritter Alpha, LP.
ritter@post.harvard.edu

Joseph Simonian is the founder and CIO of Autonomous Investment Technologies LLC, Newton, MA, USA.
jsimonian@autoinvestec.com

ABSTRACT

The Black-Litterman model is one of the most popular models in quantitative finance, with numerous theoretical and practical achievements. From the standpoint of investment theory, the Black-Litterman model allows a seamless incorporation of Bayesian statistics into the portfolio optimization process. From a practical standpoint, it provides portfolio managers with a structured approach to express subjective views, thereby freeing their investment processes from a total reliance on backward-looking historical data. In this article, the authors provide an overview of the original Black-Litterman model and its various extensions and enhancements addressing issues in real-world trading and investment management.

TOPICS

Bayesian statistics; Black-Litterman; Factor investing; Investment management; Portfolio optimization; Portfolio theory; Risk premia; Robust portfolio management; Trading; Transaction costs; Views.

KEY FINDINGS

- The authors provide an overview of the original Black-Litterman model and its various extensions and enhancements addressing issues in real-world trading and investment management.
- Particular emphasis is given to the extensions of the Black-Litterman model of significant practical relevance to today's investors and portfolio managers, including choice of priors, view generation, transaction costs, factor models, model misspecification, non-normality and multiperiod Black-Litterman portfolio optimization.

Bayesian approaches to portfolio selection have become increasingly popular over the years, primarily because they allow investors to break away from a complete reliance on historical data by incorporating their subjective views, thereby obtaining more forward-looking portfolios. Probably the most popular Bayesian approach to portfolio selection is the model developed by Black and Litterman [1990] and Black and Litterman [1991]. The Black–Litterman model (BL) provides a natural and intuitive Bayesian framework for portfolio selection, building on the familiar mean-variance optimization (MVO) model first introduced in Markowitz [1952].

A key defining feature of BL is that it gives investors the ability to incorporate their proprietary views into the inputs to the portfolio optimization process. Therefore, BL presents an advantage over the classical MVO model by freeing investors from total reliance on historical information during portfolio construction, providing more intuitively allocated portfolios and avoiding "corner solutions." An assumption underlying the BL model is that when an investor does not have a view on the market then the inputs should be consistent with the CAPM market equilibrium. As a result, an unconstrained investor with no views on the market should hold the market. In short, the BL model is a Bayesian approach that "blends" market implied expected returns with the investor's views.

The BL model is perhaps more relevant today than ever. If we consider that the most significant recent market event, the global coronavirus pandemic, was impossible to predict on the basis of time series data, it becomes readily apparent that the real-time adjustment of portfolio positioning via investor views is critical to navigating the modern market landscape. That said, given the many considerations in portfolio construction, it is not sufficient to simply form views in real-time. Rather, a rigorous and structured framework is needed to address the numerous factors that serve as building blocks to the formation of investment views. Indeed, one of the benefits of quantitative investing is that it provides a consistent and structured way of ordering market information. The BL model is a cornerstone of the standard toolkit of quantitative investing, along with statistical analysis, machine learning techniques and moderen optimization, as it provides a rigorous and scalable framework to formally incorporate subjective views into a manager's quantitative investment process.

In this article we provide an overview of the BL model and its various extensions that have been developed over the years in an effort to address issues in real-world trading and investment practice. In the first part of the article, we review the original BL model. Next, we provide a general Bayesian interpretation of the model, in which the original BL model is a special case. From here we proceed to discuss several considerations that bear on the practical implementation of the BL model. The latter include, the choice of prior, view generation, and transaction costs. The remainder of the article examines several major extensions to the original BL framework. First, we describe how the idea of the original BL model can be extended to factor models and factor investing. We next describe the application of so-called "robust optimization" methods to the BL model. Robust approaches to optimization are designed to mitigate the impact of model misspecification on portfolio construction. Our discussion of robust optimization is followed by a related discussion on various methods that have been developed to address two shortcomings of the original BL model: the assumptions of normality and linearity. In the original formulation of the BL model, the distribution of security returns are assumed to be multivariate normal, while an investor's subjective views are assumed to be linear. Both assumptions contradict empirical reality. It is well-known that many security return distributions are non-normal. For example, tail events, both positive and negative, are significantly more frequent than would be expected if security returns were normally distributed. The assumption of linearity as it pertains to investor views is also unrealistic given that many widespread investment approaches such as options strategies, have non-linear payoffs. Finally, we conclude the article with a discussion of multiperiod extensions of the BL model, in which investors can express views over different time horizons. Our ultimate goal with this article is to not only provide a historical overview of the BL model but also to describe extensions to the original model which we believe are useful to practitioners.

THE BLACK-LITTERMAN MODEL

We consider a market with n securities whose returns, $\mathbf{r}$, are multivariate normally distributed with an expected return vector, $\boldsymbol{\mu}$, and return covariance matrix, $\boldsymbol{\Sigma}$, that is

$$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})\,. \tag{1}$$

In classical portfolio theory [Markowitz, 1952], a mean-variance investor determines their portfolio holdings, $\mathbf{h} = (h_1, \ldots, h_n)'$, by solving the mean-variance optimization (MVO) problem

$$\max_{\mathbf{h}\in\mathcal{C}} \mu_p - \frac{\lambda}{2}\sigma_p^2\,, \tag{2}$$

where $\mu_p := \mathbb{E}[\mathbf{h}'\mathbf{r}] = \mathbf{h}'\boldsymbol{\mu}$ is the expected portfolio return, $\sigma_p^2 := \mathbb{V}[\mathbf{h}'\mathbf{r}] = \mathbf{h}'\boldsymbol{\Sigma}\mathbf{h}$ is the portfolio variance, $\mathcal{C}$ is the set of constraints and λ is the investor's risk aversion parameter that characterizes their trade-off between expected portfolio return and risk. Examples of constraints include long-only, trading, exposure and risk management constraints (see, Fabozzi, Kolm, Pachamanova, and Focardi [2007] and Kolm, Tütüncü, and Fabozzi [2014], for a discussion of common constraints used by portfolio managers in practice). The portfolio holdings $\mathbf{h}$ are in units of dollars, or whatever currency the investor is using. The optimal holdings of an unconstrained investor (i.e. $\mathcal{C} \equiv \mathbb{R}^n$) is given by

$$\mathbf{h}^* = \lambda^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\,.$$

In the BL model an investor can express their *subjective views* at the security level. A view might be that "the German equity market will outperform a capitalization-weighted basket of the rest of the European equity markets by 5%" [Litterman and He, 1999]. To express this view we let $\mathbf{p} = (p_1, \ldots, p_n)'$ denote the portfolio which is long one unit of the DAX index, and short a one-unit basket which holds each of the other major European indices (UKX, CAC40, AEX, etc.) in proportion to their respective aggregate market capitalizations so that $\sum_i p_i = 0$. Denoting the expected return of this portfolio by $q = 0.05$, we can express the investor's view as $\mathbb{E}[\mathbf{p}'\mathbf{r}] = q$ where $\mathbf{r}$ is the vector of security returns over the considered time horizon. If the investor has k such views

$$\mathbb{E}[\mathbf{p}_i'\mathbf{r}] = q_i, \quad \text{for} \;\; i = 1, \ldots, k\,, \tag{3}$$

the portfolios $\mathbf{p}_i$ are more conveniently arranged as rows of the $k \times n$ matrix, $\mathbf{P}$. Then the investor's views can be compactly written as

$$\mathbb{E}[\mathbf{P}\mathbf{r}] = \mathbf{q}\,, \tag{4}$$

where $\mathbf{q} := (q_1, \ldots, q_k)'$. The expectations (4) alone do not convey how confident an investor may be in their views. For this purpose, in the BL model the investor specifies a level of confidence in their views via the $k \times k$ matrix $\mathbf{\Omega}$ such that

$$\mathbf{q} = \mathbf{P}\boldsymbol{\mu} + \boldsymbol{\varepsilon}_{\mathbf{q}}, \quad \boldsymbol{\varepsilon}_{\mathbf{q}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega})\,. \tag{5}$$

Altogether, an investor's subjective views are represented as the (a) mean and (b) covariance matrix of linear combinations of the expected returns $\boldsymbol{\mu}$. This information is partial and indirect because the views are statements about portfolio returns rather than on individual security returns. Furthermore, this information is noisy, with the noise $\boldsymbol{\varepsilon}_{\mathbf{q}}$ assumed to be multivariate normally distributed, because the future is uncertain.

Black and Litterman were motivated by the principle that, in the absence of the investor's subjective views and any constraints, the mean-variance optimal portfolio holdings should be the global CAPM equilibrium portfolio, $\mathbf{h}_{\mathrm{eq}}$. Therefore, in the absence of views, the investor's model for security returns is given by (1) with

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\pi}, \mathbf{C})\,, \tag{6}$$

where $\boldsymbol{\pi}$ is the global CAPM expected return in excess of the risk-free rate and the inverse of the covariance matrix $\mathbf{C}$ represents the amount of confidence the investor has in their estimate of $\boldsymbol{\pi}$. The distributional assumption (6) is referred to as the (CAPM) *prior*.

Given the security return distribution (1), the prior (6) and the investor's views (5), the BL model yields the (posterior) expected returns

$$\boldsymbol{\mu}_{\mathrm{BL}} := (\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P} + \mathbf{C}^{-1})^{-1}(\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q} + \mathbf{C}^{-1}\boldsymbol{\pi}) \tag{7}$$

and (posterior) return covariance matrix

$$\boldsymbol{\Sigma}_{\mathrm{BL}} := \mathbb{V}[\boldsymbol{\mu}_{\mathrm{BL}}] = (\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P} + \mathbf{C}^{-1})^{-1}\,, \tag{8}$$

frequently referred to as the *Black-Litterman expected returns* and *covariance matrix*, respectively. This implies that $\mathbb{E}[\mathbf{r} \mid \mathbf{q}]$ is also given by (7) while the posterior unconditional covariance is a sum of variance due to parameter uncertainty and variance due to the randomness in $\mathbf{r}$. The optimal portfolio accounting for both types of variance is then given by

$$\mathbf{h}^*_{\mathrm{BL}} = \lambda^{-1}(\boldsymbol{\Sigma}_{\mathrm{BL}} + \boldsymbol{\Sigma})^{-1}\boldsymbol{\Sigma}_{\mathrm{BL}}(\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q} + \mathbf{C}^{-1}\boldsymbol{\pi})\,. \tag{9}$$

We refer the reader to, for example, Satchell and Scowcroft [2000], Fabozzi, Focardi, and Kolm [2006a], and Idzorek [2007] for non-Bayesian derivations of the BL model. The BL model is often described as "Bayesian" but the original authors do not elaborate directly on connections with Bayesian statistics. Kolm and Ritter [2017] provide a full Bayesian derivation of the BL model.

An important feature of the BL model is that it modifies the entire market equilibrium implied expected return vector with the investor's views. As security returns are correlated, views on just a few securities result in changes to the expected returns on all securities. This effect is stronger the more correlated the different securities are. In the absence of this adjustment of the expected return vector, the differences between the equilibrium expected return and an investor's views will be seen as an arbitrage opportunity by a mean-variance optimizer, frequently resulting in "corner portfolios" concentrated in a few holdings.

It is easy to see from the BL expected returns (7) and covariance matrix (8) that if the investor has no views or the confidence in the views is zero (i.e. $\mathbf{q} = \mathbf{0}$ or $\mathbf{\Omega} \to \infty$), then the Black-Litterman expected returns are equal to the global CAPM equilibrium returns, i.e. $\boldsymbol{\mu}_{\mathrm{BL}} \equiv \boldsymbol{\pi}$. Hence, without subjective views the unconstrained investor will hold the global CAPM portfolio.

We remark that the Black-Litterman expected returns are a "confidence" weighted linear combination of market equilibrium and investor's views, that is

$$\boldsymbol{\mu}_{\mathrm{BL}} = \mathbf{W}_{\boldsymbol{\pi}}\boldsymbol{\pi} + \mathbf{W}_{\mathbf{q}}\widetilde{\mathbf{q}}\,,$$

where $\widetilde{\mathbf{q}} := (\mathbf{P}'\mathbf{P})^{-1}\,\mathbf{P}'\mathbf{q}$ and the weighting matrices are defined by

$$\mathbf{W}_{\boldsymbol{\pi}} := (\mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{C}^{-1})^{-1}\mathbf{C}^{-1}\,, \tag{10}$$

$$\mathbf{W}_{\mathbf{q}} := (\mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{C}^{-1})^{-1}\mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P}\,, \tag{11}$$

with $\mathbf{W}_{\boldsymbol{\pi}} + \mathbf{W}_{\mathbf{q}} = \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. It can be shown that the BL expected returns (7) can also be written as $\boldsymbol{\mu}_{\mathrm{BL}} = \boldsymbol{\pi} + \mathbf{v}$, where $\mathbf{v} := \mathbf{C}\mathbf{P}'\big(\mathbf{P}\mathbf{C}\mathbf{P}' + \mathbf{\Omega}\big)^{-1}(\mathbf{q} - \mathbf{P}\boldsymbol{\pi})$ is the tilt away from the equilibrium expected excess returns. The confidence the investor has in their estimates of the market equilibrium and subjective views are represented by $\mathbf{C}^{-1}$ and $\mathbf{P}'\mathbf{\Omega}^{-1}\mathbf{P}$, respectively. In other words, when the investor has low confidence in their views, the resulting expected returns will be close to the ones implied by market equilibrium. Conversely, when the investor has higher confidence in their views, the resulting expected returns will deviate from the market equilibrium implied expected returns, thereby "tilting" away from market equilibrium.

Pros and Cons of the BL Model

As we saw in the previous section, BL relies on several key assumptions. First, the BL model assumes that security returns are normally distributed. The advantage of using this assumption within a modeling framework is that it imbues the model with mathematical tractability and transparency. The disadvantage, as we noted above is that it often leads to an inaccurate picture of security behavior. In one study, for example, 10,000 synthetic data sets were produced, each spanning approximately one century, using the GARCH (1,1), a commonly used volatility forecasting model that, like the BL model, assumes that security returns are normally distributed. In approximately one million years of simulated trading, not once in a single "GARCH century" was a market drop identified that approached the magnitude of the three largest observed in the Dow Jones Industrial Average (DJIA) during the twentieth century [Johansen and Sornette, 1998].

The assumption that investment views are (linear) portfolios of the securities is the second major assumption of the BL model. Surely, this assumption aids in simplifying the BL model, but limits its practical application. Many investment managers run investment strategies that include expressions of non-linear views. Options strategies, such as tail-risk hedging overlays, are an example of this. The original BL model is thus ill-suited for application to many types of investment strategies commonly deployed by hedge funds.

The third major simplifying assumption of the original BL model is that prior returns can be reliably derived from the CAPM market equilibrium portfolio. However, it is commonly accepted that there are a number of additional factors in addition to the market factor that are important for the derivation of prior returns. This is especially so for portfolios of multiple asset classes.

Perhaps, we may conclude that the primary advantage of the original BL framework is its simplicity. Simplicity is an important feature of any working model and indeed, one of the reasons that the BL model became so widely adopted in the first place. However, there is a trade-off between model simplicity and accuracy. In the latter respect, the original BL model is incomplete for some practical applications. The theoretical and practical extensions to the BL model we describe in the rest of the article trade some simplicity for the requisite level of accuracy required to make the resulting model a comprehensive tool in the real world.

A GENERAL FRAMEWORK: THE BLACK-LITTERMAN-BAYES MODEL

Throughout the rest of this article we discuss various extensions and developments to the BL model. As an expository tool, we will use a general Bayesian framework referred to as the *Black-Litterman-Bayes model* (BLB) introduced by Kolm and Ritter [2017]. We define a BLB model as consisting of:

1. A *parametric distribution* for the security returns $p(\mathbf{r} \,|\, \boldsymbol{\theta})$, where $\mathbf{r} := (r_1, \ldots, r_n)'$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$ denote the n security returns and m parameters, respectively;
2. A *prior distribution* $\pi(\boldsymbol{\theta})$ on the parameter space Θ;
3. A *likelihood function* $f(\mathbf{q} \,|\, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the same parameter vector appearing in the parametric statistical model above, and $\mathbf{q}$ is a k-dimensional vector supplied by the investor; and
4. A *utility function* $u(w_T)$ of final wealth, w_T, in the sense of Arrow [1971] and Pratt [1964].

In the statistics literature, items (1)-(2) are referred to as a *Bayesian statistical model.* In Bayesian statistics, all statistical inference is based on the posterior. Under such a Bayesian statistical model, decision theory shows that the optimal decision is the one maximizing posterior expected utility [Robert, 2007]. Consequently, given a BLB model, we define the associated BLB optimal portfolio holdings as the ones maximizing posterior expected utility of final wealth, that is

$$\mathbf{h}^* = \operatorname{argmax}_{\mathbf{h} \in \mathcal{C}} \mathbb{E}[u(w_0 + \mathbf{h}'\mathbf{r}) \,|\, \mathbf{q}] \tag{12}$$

where $\mathbb{E}[\,\cdot \,|\, \mathbf{q}]$ denotes the expectation with respect to the posterior predictive density for the security returns $\mathbf{r}$, $\mathbf{h} = (h_1, \ldots, h_n)'$ denotes portfolio holdings in dollars, w_0 denotes the investor's initial wealth, and $\mathcal{C}$ represents the set of constraints as before. Explicitly, the posterior predictive density of $\mathbf{r}$ is given by

$$p(\mathbf{r}\,|\,\mathbf{q}) = \int p(\mathbf{r}\,|\,\boldsymbol{\theta})p(\boldsymbol{\theta}\,|\,\mathbf{q})d\boldsymbol{\theta}\,, \quad \text{where} \tag{13}$$

$$p(\boldsymbol{\theta}\,|\,\mathbf{q}) = \frac{f(\mathbf{q}\,|\,\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int f(\mathbf{q}\,|\,\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}\,. \tag{14}$$

Suppose we are given a benchmark portfolio with holdings $\mathbf{h}_\mathrm{b}$ (such the CAPM equilibrium portfolio, $\mathbf{h}_\mathrm{eq}$). We say that the BLB prior $\pi(\boldsymbol{\theta})$ is *benchmark-optimal* if the benchmark portfolio maximizes expected utility of final wealth, where the expectation is taken with respect to the *a priori* distribution of security returns $p(\mathbf{r}) = \int p(\mathbf{r}\,|\,\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$. In particular, this is to say that

$$\mathbf{h}_\mathrm{b} = \mathrm{argmax}_{\mathbf{h}} \int u(w_0 + \mathbf{h}'\mathbf{r})p(\mathbf{r}\,|\,\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}\,. \tag{15}$$

The BLB model is not just of interest from a theoretical perspective. This Bayesian framework is quite useful for understanding and developing many of extensions of the original BL model [Kolm and Ritter, 2017]. Having said that, in general the posterior predictive density (13) and the resulting optimization problem (12) cannot be determined in closed form but need to be computed using numerical sampling and optimization techniques. However, as we will see in the next section, the BL model is special case of the BLB model for which a closed form solution is available. As we shall discuss in a later section, linear factor models represent another case in which a closed-form solution is available.

BL as a Special Case of BLB

The BL model is a special instance of the general BLB framework where $\mathbf{r}\,|\,\boldsymbol{\theta}$ is multivariate normal with mean $\boldsymbol{\theta} \equiv \boldsymbol{\mu}$, $f(\boldsymbol{\theta}\,|\,\mathbf{q}) \propto \exp\big[-\frac{1}{2}(\mathbf{P}\boldsymbol{\theta} - \mathbf{q})'\boldsymbol{\Omega}^{-1}(\mathbf{P}\boldsymbol{\theta} - \mathbf{q})\big]$ is the normal likelihood for a regression of the portfolio manager's views, the utility of final wealth is given by the exponential utility (CARA) $u(w) = -e^{-\lambda w}$, and the prior is the unique normal distribution which is benchmark-optimal with respect to the market portfolio, $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\pi}, \mathbf{C})$.

Interestingly, in the BL models there are *two* functions playing the role of likelihood functions; namely, $p(\mathbf{r} \mid \boldsymbol{\theta})$ and $f(\mathbf{q} \mid \boldsymbol{\theta})$. Similarly, the triple of random vectors $(\mathbf{r}, \mathbf{q}, \boldsymbol{\theta})$ are not pairwise independent, but $\mathbf{r} \mid \boldsymbol{\theta}$ and $\mathbf{q} \mid \boldsymbol{\theta}$ are independent. Such situations occur frequently in *Bayesian networks* (frequently also referred to as *graphical models*) [Pearl, 2014]. Bayesian networks consists of a collection of random variables whose conditional independence structure is specified by a graph. The BL model is an example of a Bayesian network with three nodes.

One of the early successes of the Black-Litterman model was that it provided a form of *regularization* of the fundamentally-unstable inversion of the historical sample covariance matrix. However, the Black-Litterman technique is not primarily meant as a regularization technique. It has that property as a by-product because its likelihood and prior tend to pull the portfolios towards known and reasonable (prior) portfolios. Importantly, with common usages of factor models of the form (21), which we discuss later in this article, regularization is strictly speaking no longer necessary because the resulting covariance matrix is already stably invertible. For example, when the mean-variance portfolio is computed using a factor-based covariance matrix, "corner portfolios" do not occur. The resulting portfolios are typically well-diversified, as long as the factor model itself is reasonable.

PRACTICAL CONSIDERATIONS WHEN USING BLB MODELS

Any application of the BL and BLB models entails consideration of a number of aspects that bear on their successful implementation. In this section, we address the key aspects that are important to any application of the BL and BLB models: (a) choice of prior, (b) view generation, and (c) transaction costs.

Choice of Prior

As discussed above, Black and Litterman were motivated by the principle that, in the absence of the investor's subjective views and any constraints, the MVO portfolio should be the capitalization-weighted CAPM benchmark. This prior is of course benchmark optimal as per the definition above, where the solution to the optimization problem (15) is given by

$$\mathbf{h}_{\text{eq}} = \lambda^{-1}(\mathbf{\Sigma} + \mathbf{C})^{-1}\boldsymbol{\pi}\,. \tag{16}$$

This follows directly from the fact that the prior is a *conjugate prior* of the normal likelihood used in the BL model, which means that the resulting posterior is of the same family (i.e. also normal). We refer to Robert [2007, Sec. 3.3] for a detailed discussion of conjugate priors in Bayesian statistics. By setting $\mathbf{C} = \tau\mathbf{\Sigma}$ with some arbitrary scalar $\tau > 0$, as did the original authors, we obtain $\boldsymbol{\pi} = \lambda(1 + \tau)\mathbf{\Sigma}\mathbf{h}_{\text{eq}}$.

There is a belief amongst some practitioners that because the BL model uses CAPM priors, the model is not useful beyond broadly diversified (global) portfolios. We remark that the formulas for the BL expected returns and covariance matrix (7) and (8) are valid for any multivariate normal prior. For active management it is common to set the mean excess return vector of the normal prior to zero [Herold, 2003; Da Silva, Lee, and Pornrojnangkool, 2009]. Such a prior leads to portable alpha implementations for active portfolio management where benchmark weights are of no interest.

At the risk of stating the obvious, one can construct market-neutral portfolios using constraints during the portfolio optimization while still using the capitalization-weighted portfolio to inform the prior. Hence BL is useful beyond long-only investing and one can still get the benefit of using the information in the market cap weights to inform the prior.

For active portfolio management such as absolute return strategies where the effective benchmark is cash, a "data-driven" approach to prior selection can be useful. Data-driven priors are very common in Bayesian statistics, where the posterior from one analysis is used as the prior for subsequent analysis. We provide an example of a data-driven prior when we describe a BL extension to linear factor models below.

View Generation

Investor views can be generated in a number of ways. Investors may simply produce them using an "all-things-considered" assessment of the information at their disposal, or they can use a quantitative framework to generate views. Traditional econometric models or modern machine learning (ML) driven frameworks are common for the latter. Given the predictive advantages that ML models have in comparison to their traditional econometric counterparts, they will be the focus of this section.[1]

At its core ML *is* prediction – a trained ML algorithm can only navigate its environment intelligently if it can make and lever accurate predictions. Perhaps more accurately, most ML approaches in common usage today are *statistical* machine learning (SML) methods, emphasizing that their training procedures correspond to fitting a statistical model to data, and their predictive mechanism consists of using the fitted statistical model to make predictions.

As ML has come into prominence over the past twenty years, the subject of how to take advantage of ML in portfolio construction has likewise been treated by many authors. The Black-Litterman model provides a natural framework to incorporate the outputs of ML models – especially statistical ML models – into the portfolio construction process. This is due to the fact that the BL model allows for the views to be input as noisy and incomplete.

Many of the early examples of Black–Litterman views were human-generated. Indeed, Exhibit 11 of Black and Litterman [1991] is a table of Goldman Sachs economists' views. When dealing with human-generated views, it can be difficult to ascertain the proper uncertainty to assign to each view. Even though the confidence in a prediction is a distinct concept from the volatility of a security's return, in the absence of any clean way to ascertain the uncertainty in the view, practitioners have sometimes simply used the covariance matrix implied by the view portfolios' holdings

$$\mathbf{\Omega} := \text{const} \cdot \mathbf{P}\mathbf{\Sigma}\mathbf{P}' \, .$$

This prescription, while convenient and trivial to implement, is not theoretically justified. What if the least volatile portfolio is also the one in which a particular economist has

[1]For a discussion of some of the advantages of ML in investment applications, see Simonian and Fabozzi [2019] and Simonian, de Prado, and Fabozzi [2018].

the least confidence in their prediction? One could argue that each portfolio has a certain intrinsic level of variance due to totally unpredictable future events, and that it is disingenuous to pretend that one's views have a lower level of error variance than the intrinsic variance due to unforeseeable future events. However, that only offers an intrinsic lower bound to the confidence in a view, not an upper bound.

Black and Litterman [1991] acknowledge that the uncertainties are not precisely known and discuss an approach in which the uncertainties are increased until the optimal portfolio becomes sufficiently balanced. Interestingly, when the views are forecasts generated by a statistical forecasting method, it becomes much more straightforward to obtain the required uncertainty parameters. Most statistical machine learning models generate not only a point prediction, but rather a full predictive density for the random variable being predicted.

Concretely, recall that in the BLB model discussed above there is a *parametric distribution* for the security returns $p(\mathbf{r} \mid \boldsymbol{\theta})$, a *prior distribution* $\pi(\boldsymbol{\theta})$ on the parameter space Θ, and a *likelihood function* $f(\mathbf{q} \mid \boldsymbol{\theta})$. In other words the single parameter vector $\boldsymbol{\theta}$ is simultaneously related to (a) the security returns $\mathbf{r}$ via the model $p(\mathbf{r} \mid \boldsymbol{\theta})$, and (b) the model describing the random variable $\mathbf{q}$ which some intelligence – human or machine – is responsible for predicting. Of course, in the classical setting $\boldsymbol{\theta}$ is the mean of $\mathbf{r}$ and $\mathbf{q}$ is defined as the return on a collection of portfolios. However, we emphasize that, in general, $\mathbf{q}$ can be any random variable which: (a) is amenable to forecasting (e.g. using ML methods), and (b) which is related to $\boldsymbol{\theta}$ by some model $f(\mathbf{q} \mid \boldsymbol{\theta})$. In particular, when the prediction of $\mathbf{q}$ is performed by an SML method, then the problem of obtaining an appropriate uncertainty matrix $\boldsymbol{\Omega}$ – typically one of the more subtle parts of applying BL in practice – becomes completely straightforward.

For instance, if the random vector $\mathbf{q}$ can be modeled as a Gaussian process (GP), then the predictive density will also be Gaussian, allowing us to use the covariance matrix of the Gaussian process prediction as $\boldsymbol{\Omega}$. Other ML methods, such as artificial neural networks (ANN) and support vector regression (SVR), naturally generate point estimates and are not usually trained to generate a full posterior predictive density. However, even in such cases, one may view such models as nonlinear regression models with Gaussian noise, and the standard errors of their predictions may be obtained via bootstrapping techniques [Fabozzi, Focardi, and Kolm, 2006b].

Transaction Costs

When using the BL model in the context of portfolio optimization, it amounts to maximizing expected utility of final wealth under the specifically formed return distribution $p(\mathbf{r} \mid \mathbf{q})$ known as the *posterior predictive density.* We recall that in the classical BL model the posterior predictive density is multivariate normal.

In the multivariate normal case, for any utility function of final wealth $u(w_T)$ that is smooth, increasing and concave, it can be shown that a maximization of expected-utility such as (12) has the same solution as the much simpler problem

$$\operatorname{argmax}_{\mathbf{h}} \; \mathbb{E}[w_T \mid \mathbf{q}] - \frac{\lambda}{2}\mathbb{V}[w_T \mid \mathbf{q}] \tag{17}$$

where w_T denotes the investor's final wealth at the investment horizon T, and $\mathbb{E}[\,\cdot \mid \mathbf{q}]$ and $\mathbb{V}[\,\cdot \mid \mathbf{q}]$ denote the expectation and variance under the posterior predictive density, respectively. Here, $\lambda > 0$ is the risk aversion which depends on the investor's utility function and initial wealth, but not on the portfolio or the returns. The optimization problem (17) is known as the *mean-variance form of the problem.* Note that this is not to say that investors have or should have quadratic utility. Rather, this mathematical fact implies that (17) achieves its optimum at the same portfolio as the harder problem (12). As Kenneth Arrow once noted, it is nonsensical for an investor to have a quadratic utility function of wealth, as at some level of wealth the investor prefers to take losses due to market moves in their investment portfolio.

The formulation (17) clarifies how transaction costs should be included into portfolio construction. First, we note that the final wealth w_T can be written

$$w_T = w_0 + \mathbf{h}'\mathbf{r} - c(\mathbf{h}) \tag{18}$$

where $c(\mathbf{h})$ denotes any costs that are incurred over the investment horizon to transition from the initial portfolio $\mathbf{h}_0$ to the target portfolio $\mathbf{h}$. Second, we observe that the risk aversion $\lambda > 0$ depends on the investor's initial wealth w_0. But once λ has been determined, the initial wealth w_0 is no longer important and can be dropped from the problem without changing the solution. Now, by inserting (18) into (17), the investor's problem becomes

$$\operatorname{argmax}_{\mathbf{h}} \ \mathbb{E}[\mathbf{h}'\mathbf{r} - c(\mathbf{h}) \,|\, \mathbf{q}] - \frac{\lambda}{2}\mathbb{V}[\mathbf{h}'\mathbf{r} - c(\mathbf{h}) \,|\, \mathbf{q}] \,. \tag{19}$$

Unless the portfolio has very high turnover, it is common in practice to make the approximation $\mathbb{V}[\mathbf{h}'\mathbf{r} - c(\mathbf{h}) \,|\, \mathbf{q}] \approx \mathbb{V}[\mathbf{h}'\mathbf{r} \,|\, \mathbf{q}]$, which is to say that most of the variance in the portfolio's return is due to market moves rather than variance in transaction costs. Consequently, a simpler, if slightly less accurate, version of the investor's problem is given by

$$\operatorname{argmax}_{\mathbf{h}} \ \mathbf{h}'\mathbb{E}[\mathbf{r} \,|\, \mathbf{q}] - \mathbb{E}[c(\mathbf{h})] - \frac{\lambda}{2}\mathbf{h}'\mathbb{V}[\mathbf{r} \,|\, \mathbf{q}]\mathbf{h} \,. \tag{20}$$

Up to the approximations discussed above, the optimization problem (20) is how any rational expected utility maximizer will perform portfolio optimization using the BL model under transaction costs.

In the classical case, $\mathbb{E}[\mathbf{r} \,|\, \mathbf{q}]$ and $\mathbb{V}[\mathbf{r} \,|\, \mathbf{q}]$ are of course the familiar Black-Litterman mean and covariance matrix given above in (7)–(8). However, we emphasize that the derivation of (20) remains valid for any general BLB model as long as the posterior predictive density of returns, $p(\mathbf{r} \,|\, \mathbf{q})$, is a member of an *elliptical family*. An elliptical density is one whose iso-probability surfaces are ellipsoids. This is the only condition one needs in order to ensure that the difficult non-linear problem of expected-utility maximization reduces to the simpler problem (20), which is a quadratic form plus a potentially-nonlinear model for expected transaction costs $\mathbb{E}[c(\mathbf{h})]$. Needless to say, this is a tremendous reduction in complexity; perhaps the most important being that one does not need to know the precise form of the investor's utility function, nor deal with its non-linearity, assuming one can determine the investor's risk aversion.

EXTENSIONS TO THE BLACK-LITTERMAN MODEL

Over time various extensions to the original BL model have been developed in an effort to bring its application closer to modern investment practice. Variants of the BL model that incorporate non-normality, nonlinear views, and model misspecification are designed to provide a more robust risk-awareness and empirical realism to the implementation of the framework. Extensions that make the BL model amenable to linear factor models and multiperiod forecasting seek to align the BL model with the actual investment processes employed by many contemporary quantitative investors.

BL for Linear Factor Models

Risk premia and factor investing have attracted considerable attention from institutional and individual investors for its simplicity, transparency and low cost. While a number of studies have addressed portfolio construction for risk premia (see, for example, Bass, Gladstone, and Ang [2017], Bergeron, Kritzman, and Sivitsky [2018], Dopfel and Lester [2018], Bender, Le Sun, and Thomas [2018], and Aliaga-Diaz, Renzi-Ricci, Daga, and Ahluwalia [2020]), most of the work in this area has focused on factor models from the frequentist's perspective.

In the BL model an investor's subjective views are expressed as portfolios of securities. Factors are not portfolios but unobservable random variables representing sources of risk that are common to all securities trading in the market. Therefore, the BL model is not directly applicable to factor views. However, Kolm and Ritter [2017] and Kolm and Ritter [2020] propose a Bayesion extension that allows for the incorporation of investor views and priors on factor risk premia. The authors show how to use the approach in a realistic empirical example with a number of well-known industry and style factors, showing their approach add value to mean-variance optimal multi-factor risk premia portfolios. We briefly describe their approach here.

Using *Arbitrage Pricing Theory* (APT) [Ross, 1976; Roll and Ross, 1980] as a starting point, we consider a market of n securities whose returns in excess of the risk-free rate at time t, denoted by $\mathbf{r}_t := (r_{t,1}, \ldots, r_{t,n})'$, are given the linear factor model

$$\mathbf{r}_t = \mathbf{X}_t \mathbf{f}_t + \boldsymbol{\varepsilon}_t\,, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{D}_t)\,, \tag{21}$$

where $\mathbf{X}_t$ is the $n \times k$ matrix of factor loadings known before time t, $\boldsymbol{\varepsilon}_t := (\varepsilon_{t,1}, \ldots, \varepsilon_{t,n})'$ is the n-dimensional random vector process of residual return, and $\mathbf{f}_t := (f_{t,1}, \ldots, f_{t,k})'$ is the k-dimensional random vector process of factors with $\mathbb{E}[\mathbf{f}_t] =: \boldsymbol{\mu}_\mathbf{f}$ and $\mathbb{V}[\mathbf{f}_t] =: \mathbf{F}$, and

$$\mathbf{D}_t := \operatorname{diag}(\sigma^2_{t,1}, \ldots, \sigma^2_{t,n}), \quad \text{where all } \sigma_{t,j} > 0, \tag{22}$$

is the $n \times n$ diagonal matrix of residual variances. The residual returns $\varepsilon_{t,i}$ are assumed to be mutually uncorrelated with each other and across time, and uncorrelated with all factors. Common examples of factors include size and value. For a size factor the loadings are typically some nonlinear transformation of the market capitalization of each stock, which are known and exogenous. The factors $f_{t,j}$ are unobservable random variables, collectively accounting for all co-movement of security returns.

Assuming $\boldsymbol{\mu}_\mathbf{f}$ and $\mathbf{F}$ are approximately constant over time, we can obtain a *factor prior* from a (Bayesian) time series model of the factor returns $\mathbf{f}_t$. For example, one possible choice is the rolling mean of a time series of the OLS estimates $\widehat{\mathbf{f}}_t = (\mathbf{X}_t'\mathbf{X}_t)^{-1}\mathbf{X}_t'\mathbf{r}_{t+1}$. More sophisticated approaches such as that of hierarchical or mixed-effects models can also be used [Gelman, Carlin, Stern, and Rubin, 2003]. We denote the prior by

$$\boldsymbol{\pi}_\mathbf{f} \sim \mathcal{N}(\boldsymbol{\xi}, \mathbf{V})\,, \tag{23}$$

where $\mathbf{V}$ is a $k \times k$ matrix represents the confidence the investor has in their estimate of $\boldsymbol{\pi}_\mathbf{f}$.

Subjective views on factor risk premia can be expressed in a similar fashion to the BL model. For instance, an investor can express views on each factor risk premium indecently, such as

$$\mathbf{q} = \boldsymbol{\mu}_\mathbf{f} + \boldsymbol{\varepsilon}_\mathbf{q}, \quad \boldsymbol{\varepsilon}_\mathbf{q} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})\,, \tag{24}$$

where $\boldsymbol{\Omega} = \text{diag}(\omega_1^2, \ldots, \omega_k^2)$ is a diagonal $k \times k$ matrix with all $\omega_j > 0$, representing the uncertainties in their views. Kolm and Ritter [2017] and Kolm and Ritter [2020] show that the (posterior) expected returns and return covariance matrix are given by

$$\boldsymbol{\mu}_{\text{BL-fac}} = \boldsymbol{\Sigma}_{\text{BL-fac}} \cdot \boldsymbol{\Sigma}^{-1}\mathbf{X}(\widetilde{\mathbf{V}}^{-1} + \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\widetilde{\mathbf{V}}^{-1}\widetilde{\boldsymbol{\xi}}\,, \tag{25}$$

$$\boldsymbol{\Sigma}_{\text{BL-fac}} = \left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\mathbf{X}(\widetilde{\mathbf{V}}^{-1} + \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\right)^{-1}, \tag{26}$$

where $\widetilde{\mathbf{V}} := (\mathbf{V}^{-1} + \boldsymbol{\Omega}^{-1})^{-1}$ and $\widetilde{\boldsymbol{\xi}} := \widetilde{\mathbf{V}}(\mathbf{V}^{-1}\boldsymbol{\xi} + \boldsymbol{\Omega}^{-1}\mathbf{q})$ are the posterior hyperparameters and

$$\boldsymbol{\Sigma} := \mathbf{D} + \mathbf{XFX} \tag{27}$$

is the covariance matrix of security returns implied by the factor model (21). In the formulas in the last paragraph, we dropped the explicit time index for purposes of readability.

Model Misspecification, Robust Optimization and BL

Model misspecification is a concern for every quantitative asset manager. As a result, for some time there has been an effort to extend the static max-min expected utility theory framework of Gilboa and Schmeidler [1989] to a dynamic environment in various areas of finance, including portfolio construction.[2] Gilboa and Schmeidler's framework provides an elegant way of modelling decision-making that incorporates investors' aversion to model uncertainty and ambiguity in the sense described by Ellsberg [1961].

The robust Black-Litterman (rBL) approach by Simonian and Davis [2011] is motivated by a desire to design a portfolio construction framework based on the BL methodology that also incorporates uncertainty about the covariance matrix of security returns, $\boldsymbol{\Sigma}$, in a simple yet coherent manner. They assume the uncertainty about $\boldsymbol{\Sigma}$ can be expressed using a distribution derived via a bootstrap procedure or an asymptotic approximation to the sampling distribution of the covariance matrix estimator.

[2] Examples of the wide variety of applications of robust methods in economics and finance include Hansen and Sargent [2001], Garlappi, Uppal, and Wang [2007], Fabozzi, Kolm, Pachamanova, and Focardi [2007], Hansen and Sargent [2008], Ben-Tal, El Ghaoui, and Nemirovski [2009], Rustem and Howe [2009], Simonian and Davis [2010], Simonian [2011], and Simonian and Davis [2011].

By taking $\mathbf{C} = \tau\boldsymbol{\Sigma}$ with some arbitrary scalar $\tau > 0$, as did Black and Litterman, there is a correspondence between each covariance matrix and the resulting the BL posterior multivariate distribution, with mean and covariance matrix given in (7) and (8). rBL characterizes a neighborhood of alternative prior distributions (6) that are also likely based on the sampling distribution of the covariance matrix as follows. For two continuous probability distributions, q and p, the Kullback-Leibler (KL) divergence, defined as

$$D_{\mathrm{KL}}(p \mid q) := \int \log\left(\frac{p(\mathbf{x})}{q(\mathbf{x})}\right) p(\mathbf{x}) d\mathbf{x}\,, \tag{28}$$

represents the information lost when one uses the q-distribution to approximate the p-distribution.

Given the multivariate normally distributed security returns (1), its KL divergence to an alternative multivariate normal distribution, $\mathcal{N}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$ is

$$D_{\mathrm{KL}}\left(\mathcal{N}_a \mid \mathcal{N}\right) = \frac{1}{2}\left(\ln\left(\frac{\det \boldsymbol{\Sigma}}{\det \boldsymbol{\Sigma}_a}\right) + \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_a\right) + \left(\boldsymbol{\mu} - \boldsymbol{\mu}_a\right)' \boldsymbol{\Sigma}\left(\boldsymbol{\mu} - \boldsymbol{\mu}_a\right) - n\right)$$

with "tr" and "det" denoting the trace and determinant, and where we constrain the alternative distribution to be consistent with the benchmark portfolio by requiring

$$\boldsymbol{\mu}_a = \lambda \boldsymbol{\Sigma}_a \mathbf{w}_{\mathrm{eq}}\,.$$

The resulting distribution can be found empirically by repeatedly drawing from the sampling distribution of $\boldsymbol{\Sigma}$. In rBL, the authors introduce another parameter to describe the confidence of the investor has in their estimate of the covariance matrix, similar to parameter τ in the original BL model that scales the investor's confidence in the prior. Define the hyperparameter δ to be the investor's confidence in the estimate of the covariance matrix, which in turn implicitly defines a value η that will be used in rBL via

$$\mathrm{P}\big(D_{\mathrm{KL}}\left(\mathcal{N}_a \mid \mathcal{N}\right) \leq \eta\big) = 1 - \delta\,,$$

where $\mathrm{P}(\cdot)$ denotes the probability measure. Thus, the investor's confidence level determines a neighborhood of alternative prior distributions that is a function of the sampling error of the covariance estimator. As the investor loses confidence in the quality of the covariance matrix estimate, the space of priors determined to be "*ex ante* plausible" expands. Acknowledging the potential imprecision of parameter estimates alters the portfolio construction process by imposing two constraints on the mean-variance objective function. The first constraint stipulates that mean excess return and covariance parameters for each security must lie within a specified confidence interval of their estimated value, explicitly recognizing the possibility of model misspecification. The second constraint expresses that a minimization over the choice of mean excess returns and covariances must be performed. Together, the constraints transform the mean-variance objective function into the max-min problem

$$\max_{\mathbf{h}} \min_{\widetilde{\boldsymbol{\Sigma}}_a} \left(\mathbf{h}' \boldsymbol{\mu}_a - \frac{\lambda}{2} \mathbf{h}' \widetilde{\boldsymbol{\Sigma}}_a \mathbf{h} \right)$$

subject to the constraint

$$\ln \left(\frac{\det \boldsymbol{\Sigma}}{\det \boldsymbol{\Sigma}_a} \right) + \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_a \right) + \left(\boldsymbol{\mu} - \boldsymbol{\mu}_a \right)' \boldsymbol{\Sigma} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_a \right) - n \leq 2\eta \, .$$

The end result is that the perturbed prior alters the entire BL process. Through robust optimization, the pivotal role of the covariance matrix in BL is brought to the fore within a Bayesian framework that explicitly incorporates the uncertainty that investors often have in their capital market assumptions.

Non-Normality and Nonlinear Views

The original work of Black and Litterman was done under the assumption the security return distribution is multivariate normal. Mathematically, this is convenient because the normal prior is a conjugate prior for the normal likelihood, and therefore the posterior is also normal with all means and covariances obtainable in closed form. Needless to say, it is of clear interest to apply BLB approaches in settings beyond the normal distribution.

Chamberlain [1983] showed that under elliptically distributed return distributions, expected utility, for *any* concave utility function, is a function of *only* the portfolio's return and variance. Elliptical distributions is a large family of distributions, including many fat-tailed distributions such as the Student's t-distributions. In the context of BLB models, this implies that if we use any form for the prior and views (including models where the unknown parameter(s) need not simply be the mean return); as long as the final posterior distribution $p(\mathbf{r}_{t+1} \,|\, \mathbf{q})$ of the security return vector is a member of the *elliptical family*, then any risk-averse expected-utility maximizer will end up maximizing the posterior expected portfolio return minus a constant multiple of the posterior portfolio variance. Naturally, in a non-normal model the investor need not arrive at the same expected return and variance formulas in the classical BL model. However, if the prior is not conjugate to whatever likelihood is chosen, then computing the posterior moments may be difficult or even impossible.

In the following subsections we discuss copula-opinion and entropy pooling, two extensions that explicitly model non-normality in security returns and allow for nonlinearities in the investor's subjective views.

Copula-Opinion Pooling

Meucci [2006a] and Meucci [2006b] propose a variant of BL model known as copula-opinion pooling (COP). Using the notation introduced for the BLB model above, $\boldsymbol{\theta}$ represents the n-dimensional vector of security returns having the prior distribution

$$\boldsymbol{\theta} \sim \pi \,. \tag{29}$$

Investors can have $k \leq n$ subjective views, which just as in the BL model are represented by a $k \times n$ matrix $\mathbf{P}$ whose i-th row determines the relative weight of each expected return in the respective view. Investors can also posit $n-k$ additional complementary linear combinations on which no view is expressed. The latter is represented in an $(n-k) \times n$ matrix denoted by $\mathbf{P}^{\perp}$. Together, the *view-adjusted coordinates* are represented by the invertible matrix

$$\overline{\mathbf{P}} := \begin{pmatrix} \mathbf{P} \\ \mathbf{P}^{\perp} \end{pmatrix}.$$

We define the n-dimensional random vector, $\mathbf{q}$, by the linear transformation of the market prior $\boldsymbol{\theta}$ as follows

$$\mathbf{q} := \overline{\mathbf{P}}\boldsymbol{\theta}. \tag{30}$$

In COP the investor's subjective views correspond to statements on the first k entries of $\mathbf{q}$. Each of the subjective views is represented in terms of a cumulative distribution function (CDF)

$$\widehat{F}_i(q) := \mathrm{P}_{\mathrm{subj}}(q_i \leq q), \quad i = 1, \ldots, k, \tag{31}$$

where the right hand side denotes the probability that the random variable q_i takes on a value less than or equal to q. In addition, the prior distribution (29) implies a distribution for each view, represented by its CDF

$$F_i(q) := \mathrm{P}_{\mathrm{prior}}(q_i \leq q), \quad i = 1, \ldots, k. \tag{32}$$

COP uses linear *opinion pooling* to derive the posterior CDF

$$\widetilde{F}_i := c_i\widehat{F}_i + (1 - c_i)\, F_i, \quad i = 1, \ldots, k, \tag{33}$$

where the pooling parameters $c_i \in [0, 1]$ represents the confidence level that investors have in each views. Accordingly, in the case when the confidence is total ($c_i = 1$), the posterior results in the investor's view; while in the total absence of confidence ($c_i = 0$), the posterior coincides with the prior CDF. The source of confidence estimates can be varied. For instance, they can be subjective inputs from portfolio managers based on the strength of their convictions or alternatively, they can be imposed exogenously by using the manager's past forecasting success (i.e. their "batting average"). Formula (33) defines the posterior marginal distribution of each view, where the copula of the posterior distribution of a given view is inherited from the copula of the market prior (29)

$$(C_1, \ldots, C_k)' := \big(F_1(q_1), \ldots, F_k(q_k)\big)' .$$

Thus the posterior joint distribution of the views becomes

$$\widetilde{\mathbf{q}} = (\widetilde{q}_1, \ldots, \widetilde{q}_k)' := \big(\widetilde{F}_1^{-1}(C_1), \ldots, \widetilde{F}_k^{-1}(C_k)\big) .$$

Finally, to determine the posterior distribution of the market $\mathbf{r} \sim p(\mathbf{r} \mid \mathbf{q})$, we apply (30) to rotate the views back into the market coordinates, obtaining

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

Simonian [2014] builds on Meucci's COP framework work by providing a method for aggregating complex opinions in the form of interconnected probability estimates using the game-theoretic solution concept known as the Shapley value.

Entropy Based Methods

In Entropy Pooling (EP), using the notation introduced for the BLB model above, Meucci [2008] assumes the "risk factors," $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$, have the prior distribution

$$\boldsymbol{\theta} \sim \pi . \tag{34}$$

For example, $\boldsymbol{\theta}$ could be security returns, latent factors in an APT model, or a combination of observed and latent factors that relate to security returns in some nonlinear way. Rather than expressing views through a likelihood function as in BLB, EP represents them by imposing constraints directly on the unknown posterior distribution of $\boldsymbol{\theta}$, denoted by $p(\boldsymbol{\theta})$, that is

$$p(\boldsymbol{\theta}) \in \mathcal{V}\,, \tag{35}$$

where $\mathcal{V}$ represents the set of these constraints. We note that the constraints can be (nonlinear) functions of the risk factors and hence quite general. Expressing the investor's subjective views in this way is one of the strengths of EP. For instance, it is straightforward to express views such as portfolios of linear and nonlinear securities, expected shortfall, and full or partial rankings of securities. The prior itself may not satisfy the constraints, $\pi \notin \mathcal{V}$. As EP does not use a likelihood function, the posterior distribution is in general not a posterior in the Bayesian sense, i.e. it cannot be determined via Bayes's theorem (14). Instead, it is computed by minimizing the KL divergence[3] (28), that is

$$p = \operatorname{argmin}_{\widetilde{p} \in \mathcal{V}} D_{\mathrm{KL}}\left(\widetilde{p} \mid \pi\right)\,. \tag{36}$$

Confidence levels for the views can also be specified in order to calibrate the mixture of the prior and posterior just as in the BL model. We point out that the computational complexity of the optimization problem (36) is determined by the functional form of the views. For a large class of problems, its dual formulation is a linearly constrained convex program of dimension equal to the number of views. Once the posterior distribution has been determined, it can be used for the purposes of portfolio optimization such as maximizing posterior expected utility of final wealth (12). We remark that while we described the parametric version of EP, it can also be implemented in a non-parametric fashion [Meucci, 2008; Meucci, Ardia, and Colasante, 2014].

Multiperiod Black-Litterman

Recently, there has been an increased interest in multiperiod portfolio optimization amongst traders and portfolio managers who make trading decisions in real-time based on time-varying forecasts of quantities such as risk, return and trading costs. For instance,

[3] Theoretically, it is of interest to note that entropy minimization generalizes Bayesian updating [Giffin, 2009]. The use of KL divergence makes robust Black-Litterman in Simonian and Davis [2011] and entropy pooling similar.

consider a systematic trader who is trading a universe of two to three thousand stocks globally. Continuously throughout the day, the trader may be producing forecasts of future returns, risk and market impact costs at different time horizons. The trader then determines trades by maximizing after-cost expected utility of portfolio value. The resulting portfolio optimization model manages a complex trade-off across the forecasts over multiple periods. Developments in portfolio trading systems and optimization have made multiperiod portfolio optimization fully possible in such real-world trading applications, even at intraday time scales (see, for example, Gârleanu and Pedersen [2013], Kolm, Tütüncü, and Fabozzi [2014], Kolm and Ritter [2015], and Boyd, Busseti, Diamond, Kahn, Koh, Nystrup, and Speth [2017]).

Consider an investor who at time t has l-step ahead forecasts of security returns up to some maximal horizon. This is referred to as an *alpha term structure*. Typically, these forecasts are increasingly uncertain as l increases. In other words, the 10-step ahead forecast typically has larger "error bars around the prediction" than the 1-step ahead forecast. This is *not* a statement that a security's return volatility is expected to increase in the future. Rather, it reflects the fact that, usually we can predict less accurately about the distant future than we can about the near future. Even for a security with constant volatility the term structure of forecast error variance can be expected to be upward-sloping.

In this context, a trade that is executed more quickly incurs more market impact, but also enjoys the benefit of exposure to the early part of the alpha term structure where the forecast is more certain. Broadly speaking, in the presence of trading costs, that forecast uncertainty increases as we look further ahead must be balanced against the market impact costs associated with trading too quickly.

Suppose one wishes to use an l-period optimization to maximize the utility of final wealth in l days. If a naive optimization method is employed, then profit made on the l-th day is viewed by the optimizer as equivalent to profit made on the first day – both increase final wealth. Hence the optimizer is willing to trade off a slightly worse portfolio on the first day for a modestly better portfolio on the l-th day. The naive approach is suboptimal because it ignores the fact that the l-step ahead forecasts have much larger forecast error variance than the one-step-ahead forecasts.

In a single-period setting, if trading costs are negligible, the BL model and its various extensions provide a coherent framework for incorporating forecast uncertainty into portfolio construction. In a setting with forecasts at multiple horizons, one needs to generalize the BL model to a multiperiod setting with trading costs [Kolm and Ritter, 2016].

The inputs to the Black–Litterman model are views, which are expected returns on given portfolios and uncertainties in those views. Suppose the current time is t, then a set of k distinct views about the future l periods from now consists of a matrix $\mathbf{P}_{t+l}$ with k rows, a vector $\mathbf{q}_{t+l}$, and a matrix $\mathbf{\Omega}_{t+l}$ where

$$\mathbf{P}_{t+l}\mathbf{r}_{t+l} = \mathbf{q}_{t+l} + \boldsymbol{\varepsilon}_{t+l}, \quad \boldsymbol{\varepsilon}_{t+l} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_{t+l}). \tag{37}$$

with $\mathbf{r}_{t+l}$ denoting the random vector of l-period-ahead security returns. In practice, often a diagonal form $\mathbf{\Omega}_{t+l} = \mathrm{diag}(\omega^2_{t+l,1}, \dots, \omega^2_{t+l,k})$ is used, indicating that the uncertainty in the views are independent. We emphasize that it is perfectly reasonable that the uncertainty terms may be statistically independent, even when the underlying view portfolios are correlated. Often the portfolio, $\mathbf{P}_{t+l}$, will be fixed as of time t, while the forecasts, $\mathbf{q}_{t+l}$, and their confidences, $\mathbf{\Omega}_{t+l}$, depend on l.

In the following, we suppress the time subscripts when they are not relevant. For any view, scaling $\mathbf{P}, \mathbf{q}$ and $\mathbf{\Omega}$ by the same constant gives an equivalent view containing the same information. Similarly, multiplying $\mathbf{P}$ and $\mathbf{q}$ both by -1 expresses an equivalent view. Hence all economically meaningful expressions, such as expressions for optimal portfolios, must be invariant under the same set of symmetries.

As discussed above, as we make predictions further into the future, our uncertainty will typically increase. This leads us to the property that elements of $\mathbf{\Omega}$ along the time axis satisfy

$$\omega_{t+l,i} < \omega_{t+l+s,i} \quad \text{for all } s > 0\,. \tag{38}$$

We emphasize that the $\omega_{t+l,i}$'s are not security return variances. There is no reason, in general, that a security's return variance should increase in the future in the manner suggested by (38). The property (38) is not a law or mathematical theorem. It is merely an expression of the fact that in most statistical forecasting models, the error variance increases the further ahead we attempt to forecast.

Assume, quite generally, that the returns on some security or some portfolio of securities are contemporaneously related to some discrete-time stochastic process Z whose value at time t is z_t. The most straightforward application is simply that z_t is the return observed in period t on some fixed portfolio of securities, but at the highest level of generality, all one needs is some statistical model connecting z_t with returns on some known portfolios. We assume that the time series z_t is amenable to forecasting via any structural time series model (such as a VAR model) that produces *l-step-ahead forecasts* for any desired integer $l > 0$ [Hamilton, 1994; Tsay, 2005]. Such models allow us to calculate the *predictive density*, $p(z_{t+l} \,|\, z_t, z_{t-1}, \ldots, z_1)$. If we assume that z_t is a Markov process, then the predictive density does not depend on the full history and can be written more simply as $p(z_{t+l} \,|\, z_t)$. The predictive distribution is often summarized by its mean and 95% quantiles. An ideal forecast $\hat{z}$ of z_{t+l} given the information set $\mathcal{F}_t$ should satisfy

$$\mathbb{E}[(z_{t+l} - \hat{z})^2 \,|\, \mathcal{F}_t] = \min_f \mathbb{E}[(z_{t+l} - f)^2 \,|\, \mathcal{F}_t] \tag{39}$$

where the minimization is over all functions of z_t. The solution is that the optimal forecast $\hat{z}$ is $\hat{z} = \mathbb{E}[z_{t+l} \,|\, z_t]$.

We define the *l-step ahead forecast error* as the random variable

$$e_t(l) := z_{t+l} - \hat{z}\,, \tag{40}$$

and the associated *forecast error variance* as

$$\sigma^2_{t+l\,|\,t} := \mathbb{V}[e_t(l) \,|\, \mathcal{F}_t]\,. \tag{41}$$

The truly fundamental object (in the sense that it contains all of the information we have about various periods in the future) is the collection of predictive densities, $\{p(z_{t+l} \mid z_t) : l = 1, 2, 3, \ldots\}$, which we call *the term structure of information.* As long as the variable being forecasted, z_t, is identifiable as the return on some portfolio of securities, referring to (37), we may set a row of $\mathbf{P}$ to be this portfolio, and then identify the Black-Litterman quantities q, ω in terms of forecasts produced by our statistical model

$$q_{t+l} := \mathbb{E}[z_{t+l} \mid z_t], \quad \omega_{t+l}^2 := \mathbb{V}[e_t(n) \mid \mathcal{F}_t] = \sigma_{t+l\,|\,t}^2. \tag{42}$$

The formulas in (42) are the basic relations connecting the uncertainty variances ω_{t+l}^2, which are inputs to the Black–Litterman procedure, to the output of multi-step-ahead statistical forecast procedures.

Next, we provide a concrete example. Suppose we have two securities with returns x_t, y_t in period t. Suppose for some constant β, that

$$z_t := y_t - \beta x_t$$

is the stationary return on a portfolio which is short β units of the first security and long one unit of the second secucrity. We define a matrix $\mathbf{P}$ whose one single row is the portfolio

$$\mathbf{P} = [-\beta \quad 1]\,. \tag{43}$$

Then by construction

$$\mathbf{P}\mathbf{r}_{t+l} = y_{t+l} - \beta x_{t+l} \equiv z_{t+l}\,.$$

We take the q_{t+l} and ω_{t+l} for all l to be given by (42), and assume for simplicity that z_t follows the mean-zero AR(1) model

$$z_t - \phi z_{t-1} = a_t \text{ with } |\phi| < 1 \text{ and } a_t \sim \mathcal{N}(0, \sigma_a^2).$$

Then, a standard calculation shows that

$$\hat{z}(t+l) = \phi^n z_t$$
$$\mathbb{V}[e_t(n)] = (1 + \phi^2 + \ldots + \phi^{2(n-1)})\sigma_a^2.$$

Note that as $l \to \infty$, we have $\hat{z}(t+l) \to 0$ which is the long-run mean, and moreover the variance of forecast error converges to $(1-\phi^2)^{-1}\sigma_a^2$, which is the variance of z_t.

In the following, we denote by δ the first-difference operator for time series, so for any time series $\{x_t : t = 1, \ldots, T\}$, we have $\delta x_t := x_t - x_{t-1}$. Let $\mathbf{r}_t$ denote the n-dimensional vector of security returns at times $t = 1, \ldots, T$. We are interested in our wealth at time T given by

$$w_T = w_0 + \sum_{t=1}^{T} \delta w_t\,, \tag{44}$$

where $\delta w_t = w_t - w_{t-1}$ denotes the change in wealth associated to period t. With $\mathbf{h}_t$ denoting the portfolio holdings at the beginning of the period at time t, then

$$\delta w_t = \mathbf{h}_t' \mathbf{r}_t - c(\mathbf{h}_t, \mathbf{h}_{t-1}) \tag{45}$$

where $c(\mathbf{h}_t, \mathbf{h}_{t-1})$ includes all forms of cost that could potentially cause trading profit to differ from its idealized form $\mathbf{h}_t'\mathbf{r}_t$.

The full collection of security returns $\mathbf{r} := (\mathbf{r}_1', \ldots, \mathbf{r}_T')'$ is then an nT-dimensional random column vector. This vector is expressed in *stacked notation*, a way of representing a T-period problem of n securities as a one-period problem of nT securities. We assume all column vectors are stacked vertically. All matrices become block matrices, with T separate $n \times n$ blocks stacked along the diagonal. Any blocks which are not one of the $n \times n$ blocks arranged along the diagonal represent relationships that connect one time period with another, such as serial autocorrelation. Intuitively, this amounts to treating IBM shares today and IBM shares tomorrow as two separate entities that we might consider trading; a portfolio must specify holdings for both "versions" of IBM.

Let $\mathbf{h} := (\mathbf{h}_1', \ldots, \mathbf{h}_T')'$ denote the full sequence of portfolios – this is the decision variable in multiperiod portfolio optimization. Final wealth w_T is implicitly a function of $\mathbf{h}$. If the probability law of $\mathbf{r}$ has an elliptical distribution, then to optimize

$$\mathbf{h}^* = \operatorname{argmax}_{\mathbf{h}} \mathbb{E}[u(w_T)]$$

one may equivalently optimize

$$\begin{aligned} \mathbf{h}^* &= \operatorname{argmax}_{\mathbf{h}} \left\{ \mathbb{E}[w_T] - \frac{\lambda}{2} \mathbb{V}[w_T] \right\} \qquad (46) \\ &= \operatorname{argmax}_{\mathbf{h}} \left\{ \mathbb{E}[\mathbf{h}'\mathbf{r} - c(\mathbf{h})] - \frac{\lambda}{2} \mathbb{V}[\mathbf{h}'\mathbf{r} - c(\mathbf{h})] \right\} \end{aligned}$$

with $c(\mathbf{h}) = \sum_{t=1}^{T} c(\mathbf{h}_t, \mathbf{h}_{t-1})$ and where we used formulas (44) and (45).

In stacked notation, the more complicated multiperiod problem resembles a one-period problem. In particular, there is a multiperiod analogue of (37). A separate view at each time-step is expressed as

$$\mathbf{P}_t \mathbf{r}_t = \mathbf{q}_t + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_t)$. The number of views at step t is denoted by k_t. We allow the possibility that $k_t \neq k_s$ for $t \neq s$, but the more familiar situation is that we have a single set of view portfolios, but the uncertainty increases as we look further into the future. In this more familiar situation, the $k_t \times N$ matrices $\mathbf{P}_t$ for various t are all the same matrix. By stacking the views, we define

$$\mathbf{P} := \begin{pmatrix} \mathbf{P}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}_T \end{pmatrix}, \qquad \mathbf{q} := \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \vdots \\ \mathbf{q}_T \end{pmatrix}$$

so that $\mathbf{q}$ is an nT-dimensional vector and the dimension of the $\mathbf{P}$ matrix is $(\sum_t k_t) \times nT$.

In *the absence of trading costs*, it is straightforward to extend the BL methodology to a multiperiod setting. We will do this first, and return to the problem with trading costs. Assuming the covariance matrix of security returns, $\boldsymbol{\Sigma}$, is known and the prior is Gaussian as in the original BL model, it is easy to see that the posterior distribution $p(\boldsymbol{\theta} \,|\, \mathbf{q})$ has the mean and covariance matrix given formulas (7) and (8) where all matrices and vectors are replaced with their stacked equivalents. Similarly, this implies that the mean-variance optimal multiperiod portfolio is given by (9).

In *the presence of trading costs* the full problem is given by

$$\operatorname{argmax}_{\mathbf{h}} \left\{ \mathbb{E}[\mathbf{h}'\mathbf{r} - c(\mathbf{h}) \,|\, \mathbf{q}] - \frac{\lambda}{2} \mathbb{V}[\mathbf{h}'\mathbf{r} - c(\mathbf{h}) \,|\, \mathbf{q}] \right\} \tag{47}$$

where $\mathbb{E}$ and $\mathbb{V}$ are computed using the posterior predictive density $p(\mathbf{r} \,|\, \mathbf{q})$. We take the simplifying assumption that costs are certain, and expected costs are not modified by views (so $\mathbb{E}[c(\mathbf{h})]$ is not a function of $\mathbf{P}, \mathbf{q}$). We further assume no serial dependence among security returns. Then, the objective function of (47) becomes

$$\sum_{t=1}^{T} \left\{ \mathbf{h}_t'\mathbb{E}[\mathbf{r}_t \,|\, \mathbf{q}_t] - \frac{\lambda}{2} \mathbf{h}_t'\mathbb{V}[\mathbf{r}_t \,|\, \mathbf{q}_t]\mathbf{h}_t - c(\delta\mathbf{h}_t) \right\} , \tag{48}$$

where

$$\delta\mathbf{h}_t := \mathbf{h}_t - \mathbf{h}_{t-1}, \quad \mathbf{h}_t := \mathbf{h}_0 + \delta\mathbf{h}_1 + \cdots + \delta\mathbf{h}_t \,.$$

Here $\delta\mathbf{h}_t$ is called a *trade list* or *order wave*.

Typically, the expected cost of a portfolio transition is separable over securities, so $c(\delta\mathbf{h}_t) = \sum_{i=1}^{N} c_i(\delta\mathbf{h}_{t,i})$. For a single security, the function c_i typically includes both bid-offer spread cost, which depends on the absolute value of the trade size, as well as predicted market impact which is usually the square or $3/2$ power of the absolute trade size, with a coefficient dependent upon the expected volume and volatility. Such cost functions are convex, but can be pointwise non-differentiable. Consequently, the multiperiod problem (48) will be convex in which the only non-differentiable terms are separable across securities and across time. This implies that the globally optimal solution can be found by convex optimization techniques.

CONCLUSION

The Black-Litterman model is one of the most successful models in finance. Its appeal stems from its roots in Bayesian statistics, which provides the mathematical tools to combine investors' forward-looking subjective views with a market-based prior. The latter quality has proven attractive to investment practitioners who must constantly confront the twin challenges of limited data sets and rapidly changing market information. Since the introduction of the Black-Litterman model in the early 1990s, there has been significant advancements in statistical inference techniques, often driven by advances in machine learning and the need to analyze large complex datasets in real-time. Moreover, many practical problems in finance require inference in "data scarce" situations where the number of parameters may greatly exceed the number of observations. The Black-Litterman model is thus as relevant as ever. In this article, the authors have provided an overview of the original Black-Litterman model, its major implementation considerations and some of the important extensions to the original framework. The authors hope that by providing this exploration of the Black-Litterman model, academics and practitioners will be inspired to further develop it in response to the evolving needs of modern investment practice.

REFERENCES

Aliaga-Diaz, Roger, Giulio Renzi-Ricci, Ankul Daga, and Harshdeep Ahluwalia (2020). "Portfolio Optimization with Active, Passive, and Factors: Removing the Ad Hoc Step". In: *The Journal of Portfolio Management* 46.4, pp. 39–51.

Arrow, Kenneth J (1971). *Essays in the Theory of Risk-Bearing*. North-Holland Amsterdam.

Bass, Robert, Scott Gladstone, and Andrew Ang (2017). "Total Portfolio Factor, Not Just Asset, Allocation". In: *The Journal of Portfolio Management* 43.5, pp. 38–53.

Ben-Tal, Aharon, Laurent El Ghaoui, and Arkadi Nemirovski (2009). *Robust Optimization*. Vol. 28. Princeton University Press.

Bender, Jennifer, Jerry Le Sun, and Ric Thomas (2018). "Asset Allocation vs. Factor Allocation—Can We Build a Unified Method?" In: *The Journal of Portfolio Management* 45.2, pp. 9–22.

Bergeron, Alain, Mark Kritzman, and Gleb Sivitsky (2018). "Asset Allocation and Factor Investing: An Integrated Approach". In: *The Journal of Portfolio Management* 44.4, pp. 32–38.

Black, Fischer and Robert B. Litterman (1990). *Asset Allocation: Combining Investor Views with Market Equilibrium.* Tech. rep. Goldman Sachs & Co.

— (1991). *Global Asset Allocation with Equities, Bonds and Currencies*. Tech. rep. Goldman Sachs & Co.

Boyd, Stephen, Enzo Busseti, Steven Diamond, Ronald N. Kahn, Kwangmoo Koh, Peter Nystrup, and Jan Speth (2017). *Multi-Period Trading via Convex Optimization.*

Chamberlain, Gary (1983). "A Characterization of the Distributions That Imply Mean—Variance Utility Functions". In: *Journal of Economic Theory* 29.1, pp. 185–201.

Da Silva, Alexandre S., Wai Lee, and Bobby Pornrojnangkool (2009). "The Black–Litterman Model for Active Portfolio Management". In: *The Journal of Portfolio Management* 35.2, pp. 61–70.

Dopfel, Frederick E. and Ashley Lester (2018). "Optimal Blending of Smart Beta and Multifactor Portfolios". In: *The Journal of Portfolio Management* 44.4, pp. 93–105.

Ellsberg, Daniel (1961). "Risk, Ambiguity, and the Savage Axioms". In: *The Quarterly Journal of Economics*, pp. 643–669.

Fabozzi, Frank J., Sergio M. Focardi, and Petter N. Kolm (2006a). *Financial Modeling of the Equity Market: From CAPM to Cointegration.* John Wiley & Sons.

— (2006b). "Incorporating Trading Strategies in the Black-Litterman Framework". In: *The Journal of Trading* 1.2, pp. 28–37.

Fabozzi, Frank J., Petter N. Kolm, Dessislava A. Pachamanova, and Sergio M. Focardi (2007). *Robust Portfolio Optimization and Management.* John Wiley & Sons.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang (2007). “Portfolio Selection with Parameter and Model Uncertainty: A Multi-prior Approach”. In: *The Review of Financial Studies* 20.1, pp. 41–81.

Gârleanu, Nicolae and Lasse Heje Pedersen (2013). “Dynamic Trading with Predictable Returns and Transaction Costs”. In: *The Journal of Finance* 68.6, pp. 2309–2340.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin (2003). *Bayesian Data Analysis*. 2nd. Taylor & Francis.

Giffin, Adom (2009). “Maximum Entropy: The Universal Method for Inference”. In: *arXiv preprint arXiv:0901.2987*.

Gilboa, Itzhak and David Schmeidler (1989). “Maxmin Expected Utility with a Non-Unique Prior”. In: *Journal of Mathematical Economics* 18.2, pp. 141–153.

Hamilton, James Douglas (1994). *Time Series Analysis*. Princeton University Press.

Hansen, Lars Peter and Thomas J. Sargent (2001). “Robust Control and Model Uncertainty”. In: *American Economic Review* 91.2, pp. 60–66.

— (2008). *Robustness*. Princeton University Press.

Herold, Ulf (2003). “Portfolio Construction with Qualitative Forecasts”. In: *The Journal of Portfolio Management* 30.1, pp. 61–72.

Idzorek, Thomas (2007). “A Step-by-Step Guide to the Black-Litterman Model: Incorporating User-Specified Confidence Levels”. In: *Forecasting Expected Returns in the Financial Markets*. Ed. by Stephen Satchell. Elsevier, pp. 17–38.

Johansen, Anders and Didier Sornette (1998). “Stock Market Crashes are Outliers”. In: *The European Physical Journal B-Condensed Matter and Complex Systems* 1.2, pp. 141–143.

Kolm, Petter N. and Gordon Ritter (2015). “Multiperiod Portfolio Selection and Bayesian Dynamic Models”. In: *Risk* 28.3, pp. 50–54.

— (2016). “Multiperiod Black-Litterman Portfolio Optimization”. Unpublished manuscript.

— (2017). “On the Bayesian Interpretation of Black-Litterman”. In: *European Journal of Operational Research* 258.2, pp. 564–572.

Kolm, Petter N. and Gordon Ritter (2020). “Factor Investing with Black–Litterman–Bayes: Incorporating Factor Views and Priors in Portfolio Construction”. In: *The Journal of Portfolio Management* 47.2, pp. 113–126.

Kolm, Petter N., Reha Tütüncü, and Frank J. Fabozzi (2014). “60 Years of Portfolio Optimization: Practical Challenges and Current Trends”. In: *European Journal of Operational Research* 234.2, pp. 356–371.

Litterman, Robert and Guangliang He (1999). *The Intuition Behind Black-Litterman Model Portfolios*. Tech. rep. Goldman Sachs.

Markowitz, Harry (1952). “Portfolio Selection”. In: *The Journal of Finance* 7.1, pp. 77–91.

Meucci, Attilio (2006a). “Beyond Black-Litterman in Practice: A Five-Step Recipe to Input Views on Non-Normal Markets”. In: *Available at SSRN 872577*.

— (2006b). “Beyond Black-Litterman: Views on Non-Normal Markets”. In: *Risk* 19, pp. 87–92.

— (2008). “Fully Flexible Views: Theory and Practice”. In: *Risk* 21.10, pp. 97–102.

Meucci, Attilio, David Ardia, and Marcello Colasante (2014). “Portfolio Construction and Systematic Trading with Factor Entropy Pooling”. In: *Risk* 27.5, pp. 56–61.

Pearl, Judea (2014). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Pratt, John W. (1964). “Risk Aversion in the Small and in the Large”. In: *Econometrica: Journal of the Econometric Society*, pp. 122–136.

Robert, Christian (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer Science & Business Media.

Roll, Richard and Stephen A. Ross (1980). “An Empirical Investigation of the Arbitrage Pricing Theory”. In: *The Journal of Finance* 35.5, pp. 1073–1103.

Ross, Stephen A. (1976). “The Arbitrage Theory of Capital Asset Pricing”. In: *Journal of Economic Theory* 13.3, pp. 341–360.

Rustem, Berc and Melendres Howe (2009). *Algorithms for Worst-Case Design and Applications to Risk Management*. Princeton University Press.

Satchell, Stephen and Alan Scowcroft (2000). “A Demystification of the Black–Litterman Model: Managing Quantitative and Traditional Portfolio Construction”. In: *Journal of Asset Management* 1.2, pp. 138–150.

Simonian, Joseph (2011). “A Bayesian Approach to Building Robust Structural Credit Default Models”. In: *Applied Economics Letters* 18.14, pp. 1397–1400.

— (2014). “Copula-Opinion Pooling with Complex Opinions”. In: *Quantitative Finance* 14.6, pp. 941–946.

Simonian, Joseph and Josh Davis (2010). “Robust Value-at-Risk: An Information-Theoretic Approach”. In: *Applied Economics Letters* 17.16, pp. 1551–1553.

— (2011). “Incorporating Uncertainty into the Black–Litterman Portfolio Selection Model”. In: *Applied Economics Letters* 18.17, pp. 1719–1722.

Simonian, Joseph, Marcos López de Prado, and Frank J. Fabozzi (2018). “Invited Editorial Comment: Order from Chaos - How Data Science Is Revolutionizing Investment Practice”. In: *The Journal of Portfolio Management* 45.1, pp. 1–4.

Simonian, Joseph and Frank J. Fabozzi (2019). “Triumph of the Empiricists: The Birth of Financial Data Science”. In: *The Journal of Financial Data Science* 1.1, pp. 10–13.

Tsay, Ruey S. (2005). *Analysis of Financial Time Series*. Vol. 543. John Wiley & Sons.