



# 60 Years of portfolio optimization: Practical challenges and current trends



Petter N. Kolm<sup>a,\*</sup>, Reha Tütüncü<sup>b</sup>, Frank J. Fabozzi<sup>c</sup>

<sup>a</sup> Courant Institute, New York University, 251 Mercer St., New York, NY 10012, United States

<sup>b</sup> Goldman Sachs Asset Management, 200 West St., New York, NY 10282, United States

<sup>c</sup> EDHEC Business School, 393, Promenade des Anglais BP3116, 06202 Nice Cedex 3, France

## ARTICLE INFO

### Article history:

Available online 1 November 2013

### Keywords:

Black–Litterman  
Estimation errors  
Mean–variance optimization  
Multi-period optimization  
Portfolio constraints  
Portfolio optimization

## ABSTRACT

The concepts of portfolio optimization and diversification have been instrumental in the development and understanding of financial markets and financial decision making. In light of the 60 year anniversary of Harry Markowitz's paper "Portfolio Selection," we review some of the approaches developed to address the challenges encountered when using portfolio optimization in practice, including the inclusion of transaction costs, portfolio management constraints, and the sensitivity to the estimates of expected returns and covariances. In addition, we selectively highlight some of the new trends and developments in the area such as diversification methods, risk-parity portfolios, the mixing of different sources of alpha, and practical multi-period portfolio optimization.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

The concepts of portfolio optimization and diversification have been instrumental in the development and understanding of financial markets and financial decision making. The major breakthrough came in 1952 with the publication of Harry Markowitz's theory of portfolio selection (Markowitz, 1952). The theory, popularly referred to as *modern portfolio theory*, provided an answer to the fundamental question: How should an investor allocate funds among the possible investment choices? First, Markowitz quantified return and risk of a security, using the statistical measures of its expected return and standard deviation. Second, Markowitz suggested that investors should consider return and risk together, and determine the allocation of funds among investment alternatives on the basis of their return-risk trade-off. Before Markowitz's seminal article, the finance literature had treated the interplay between return and risk in an ad hoc fashion.

The idea that sound financial decision-making is a quantitative trade-off between return and risk was revolutionary for two reasons. First, it posited that one could make a quantitative evaluation of portfolio return and risk jointly by considering security returns and their co-movements. An important principle at work here is that of portfolio diversification. It is based on the idea that a portfolio's riskiness depends on the correlations of its constituents, not only on the average riskiness of its separate holdings. This concept

was foreign to classical financial analysis, which revolved around the notion of the value of single investments, that is, the belief that investors should invest in those assets that offer the highest future value given their current price. Second, it formulated the financial decision-making process as an optimization problem. In particular, the so-called mean–variance optimization (MVO) problem formulated by Markowitz suggests that among the infinite number of portfolios that achieve a particular return objective, the investor should choose the portfolio that has the smallest variance. All other portfolios are "inefficient" because they have a higher variance and, therefore, higher risk.

Markowitz's work has had a major impact on academic research and the financial industry as a whole. Some internet searches we did as of the writing of this paper revealed the following numbers:

- 19,016 articles in Google Scholar cite Markowitz's original paper "Portfolio Selection".
- When searching for "modern portfolio theory" we obtained:  
About 590,000 hits in Google.  
531 YouTube videos.  
217 books on Amazon.  
Many thousands of tweets on Twitter.

MVO is used both for constructing portfolios of individual assets (*asset level*) and for asset allocation (*asset class level*). While in this paper we focus on the former application, the majority of the techniques we discuss are applicable to asset allocation, optimization on the asset class level is often considered easier than on the asset level, primarily because of the small number of asset classes. Interestingly, today more than 60 years later, risk-return optimization

\* Corresponding author. Tel.: +1 (212) 998 4855.

E-mail addresses: [petter.kolm@nyu.edu](mailto:petter.kolm@nyu.edu) (P.N. Kolm), [reha.tutuncu@gs.com](mailto:reha.tutuncu@gs.com) (R. Tütüncü), [frank.fabozzi@edhec.edu](mailto:frank.fabozzi@edhec.edu) (F.J. Fabozzi).

at the asset level is still primarily done only at the larger and/or more quantitatively oriented firms. However, with the availability of optimization tools customized for portfolio and risk management more and more investment managers are using some form of risk-return optimization as part of their portfolio construction process. A major reason for the surprisingly slow adaptation by investment managers to apply quantitative risk-return optimization is that they have observed that directly “out-of-the-box” portfolio optimization tends to be unreliable in practice. Specifically, risk-return optimization can be very sensitive to changes in the inputs, especially when the return and risk estimates are not well aligned or when the problem formulation uses multiple, interacting constraints. As a result, many practitioners consider the output of risk-return optimization to be opaque, unstable, and/or unintuitive.

Estimation errors in the forecasts significantly impact the resulting portfolio weights. For example, it is well-known that in practical applications equally weighted portfolios often outperform mean-variance portfolios (DeMiguel, Garlappi, & Uppal, 2009; Jobson & Korkie, 1981; Jorion, 1985), mean-variance portfolios are not necessarily well-diversified (Green & Hollifield, 1992), portfolio optimizers are often “error maximizers” (Michaud, 1998), and mean-variance optimization can produce extreme or non-intuitive weights for some of the assets in the portfolio (Black & Litterman, 1991, 1992). Such examples, however, are not necessarily a sign that the theory of risk-return optimization is flawed. Rather, it means that the classical framework has to be modified *when used in practice* in order to achieve reliability, stability, and robustness with respect to model and estimation errors. We will review some of the common approaches for this purpose in this paper.

Our intention with this article is not to provide a survey of MVO, its extensions and related areas. Some surveys include Steinbach (2001), Rubinstein (2002), Fabozzi, Kolm, Pachamanova, and Focardi (2007), and Markowitz (2014). Admittedly, there are many important contributions and works that we do not cover due to space constraints. The main goal with this article is twofold.

First, we address some of the key aspects related to using portfolio optimization in practice. The inclusion of transaction costs in the portfolio selection problem may present a challenge to the portfolio manager, but is an important practical consideration. We discuss a standard approach on how to extend traditional asset allocation models to incorporate transaction costs. In practice, it is common to amend the mean-variance framework with various types of constraints that take specific investment guidelines and institutional features into account. We discuss the use of various categories of constraints in portfolio construction and methods that quantify their impact on the portfolios generated. One of the main criticisms of the MVO approach focuses on its dependence on estimated parameters; specifically, expected returns and covariances, and its sensitivity to errors in these estimates. We outline various approaches that exist in the literature to mitigate the impact of estimation errors, including Bayesian methods, the Black-Litterman approach, and robust optimization techniques.

Second, we selectively highlight some of the new trends and developments in MVO and its related areas. Due to space constraints, we cannot survey all new trends in this area. While admittedly our choice is subjective, it is based on what we believe are as some of the important developments in this area for the use of MVO and its extensions in practice. In particular, we discuss the recent focus on diversification methods and provide a summary of the developments related to risk-parity portfolios. We also provide a formalization of the problem of mixing different sources of alpha and address some of the challenges that arise from these formulations. Finally, we outline some of the recent literature on practical multi-period portfolio optimization.

The paper is organized as follows. In Section 2, we review classical MVO. In Section 3, we discuss some of the most common ways on how to address the challenges one encounters when implementing MVO in practice. In Section 4, we highlight some interesting new directions and trends in MVO and related areas.

## 2. Mean-variance optimization

We consider an investment universe of  $n$  assets  $S_1, S_2, \dots, S_n$  with uncertain future returns  $r_1, r_2, \dots, r_n$ . We denote by  $r = [r_1, \dots, r_n]^T$  the vector of these returns. A portfolio is represented by the  $n$ -dimensional vector  $\omega = [\omega_1, \dots, \omega_n]^T$  where  $\omega_i$  denotes the proportion of the total funds invested in security  $i$ . The (uncertain) return of the portfolio,  $r_p$ , depends linearly on the weights

$$r_p(\omega) = \omega_1 r_1 + \dots + \omega_n r_n = \omega^T r$$

We denote by  $\sigma_i$  the standard deviation of  $r_i$ ,  $\rho_{ij}$  denote the correlation coefficient of the returns of assets  $S_i$  and  $S_j$  (for  $i \neq j$ ), and  $\Sigma$  the (symmetric)  $n \times n$  covariance matrix of the returns of all the assets, i.e.

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$$

where  $\sigma_{ii} = \sigma_i^2$  and  $\sigma_{ij} = \sigma_{ji} = \rho_{ij}\sigma_i\sigma_j$  (for  $i \neq j$ ). All valid covariance matrices are positive semidefinite matrices (i.e.  $\omega^T \Sigma \omega \geq 0$  for all  $\omega$ ), or equivalently, all of their eigenvalues are nonnegative. In this paper, we assume that  $\Sigma$  satisfies the stronger property of positive definiteness, namely that  $\omega^T \Sigma \omega > 0$  for all  $\omega \neq 0$ . This is equivalent to assuming that none of the assets  $S_1, S_2, \dots, S_n$  can be perfectly replicated by a combination of the remaining assets. Positive definiteness assumption ensures that  $\Sigma$  is an invertible matrix. For a given portfolio  $\omega$ , we can compute the variance and the standard deviation of the portfolio return as

$$\begin{aligned} V(\omega) &= \omega^T \Sigma \omega \\ \sigma(\omega) &= \sqrt{\omega^T \Sigma \omega} \end{aligned}$$

The standard deviation of the portfolio return  $\sigma(\omega)$ , also referred to as portfolio volatility, is frequently used as a measure of risk of the portfolio  $\omega$ .

We let  $\Omega$ , a subset of  $\mathbb{R}^n$ , denote the set of permissible portfolios. In particular,  $\omega \in \Omega$  means that the portfolio weights have to satisfy the constraints we impose upon our portfolio.

We represent the expected returns of the securities by

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}$$

where  $\mu_i = E(r_i)$  for all  $i = 1, \dots, n$ .

Using this notation, the MVO problem takes the form

$$\max_{\omega \in \Omega} \mu^T \omega - \lambda \cdot \omega^T \Sigma \omega$$

where  $\lambda$  is an investor specific risk aversion parameter that determines the trade-off between expected portfolio return and portfolio risk.

Alternative formulations of the MVO problem are obtained by either maximizing the expected return subject to an upper limit on the portfolio variance, or by minimizing the portfolio variance subject to a lower limit on the expected return, i.e.

$$\begin{aligned} \max_{\omega \in \Omega} \mu^\top \omega \\ \omega^\top \Sigma \omega \leq \sigma_{\max}^2 \\ \text{or} \\ \min_{\omega \in \Omega} \omega^\top \Sigma \omega \\ \mu^\top \omega \geq R_{\min} \end{aligned}$$

If the constraint set only includes linear equality and inequality constraints then the MVO is a *quadratic program* (QP) and can be solved by using standard numerical optimization software. Modern portfolio optimization software can also deal with nonlinear constraints, such as risk limits or risk contribution limits on groups of securities, as well as constraints with discrete elements such as number of holdings and/or trades constraints. Such formulations are typically solved by using software that has conic optimization and integer optimization capabilities.<sup>1</sup>

### 3. Advances in using MVO in practice

Markowitz's seminal paper on portfolio selection undoubtedly has had a major impact not only on academic research but on the financial industry as a whole. It changed the focus of investment analysis away from individual security selection toward the concept of diversification and the impact of individual securities on a portfolio's risk-return characteristics. In the MVO framework, efficient portfolios are formed by choosing an asset based upon its interaction with other assets in the portfolio as well as on its contribution to the overall portfolio, and not on the basis of its stand-alone performance.

Despite the simplicity and intuitive appeal of portfolio construction using modern portfolio theory, it took many years until portfolio managers started using portfolio optimization to manage real money. In real world applications there are many concerns associated with its use, and portfolio optimization is still considered by many practitioners to be impractical to apply. As we argued in the introduction, some of these concerns are related to the often unintuitive relationship between the inputs and outputs of MVO as well as the sensitivity of the optimal portfolio allocation to changes in the inputs, inputs that are inherently difficult to estimate. Some practitioners consider the notion of optimality offered by modern portfolio theory to be ill-defined, or of "false precision", in the noisy and non-stationary system formed by financial security returns. The perceived robustness and intuitiveness of simpler, rules-based portfolio construction approaches may be preferred to the mathematical rigor of optimal portfolio selection.<sup>2</sup>

In recognition of these concerns, the original approach proposed by Markowitz only serves as a starting point and the classical mean-variance framework is often extended in several different directions for portfolio management in practice. These extensions include, but are not limited to, the following:

- The inclusion of transaction costs (such as market impact costs) and tax effects.<sup>3</sup>

- The addition of various types of constraints that take specific investment guidelines and institutional features into account.
- Modeling and quantification of the impact of estimation errors in risk and return forecasts on the portfolios via Bayesian techniques, stochastic optimization, or robust optimization approaches.
- Multi-period extensions of MVO to incorporate intertemporal effects such as hedging needs, changing market conditions, market impact costs, and alpha decay.

We discuss the first three topics in this section and defer the discussion of the fourth topic to Section 4.

#### 3.1. Incorporating transaction costs into MVO

The objective of the portfolio allocation process is to find an optimal trade-off between return and risk. Traditionally, this was done independently of trading cost considerations, as the control and management of trading costs were handled separately by the trading desk. This suboptimal approach would often lead to target portfolio holdings that would incur large trading costs, in some cases having a severe impact on realized risk-adjusted returns. By directly incorporating transaction costs into the portfolio allocation process, resulting portfolios are more cost effective and show improvement in terms of realized risk-adjusted returns. Transaction costs make the portfolio optimization problem difficult to solve when the number of assets is greater than two, especially in a dynamic setting. For a recent example, see [Brown and Smith \(2011\)](#).

Transaction costs consist of direct costs, such as commissions and taxes, bid-ask spread, and indirect costs, such as slippage. Slippage is the difference between the price prevailing at the time the trade is anticipated,  $t_0$ , and the volume weighted average price over the time period,  $[t_0, t_0 + T]$ , over which it executes. Somewhat simplistically, slippage is due to (a) random price changes in the securities that occur in  $[t_0, t_0 + T]$ , and (b) "market impact costs," i.e. price changes incurred because of the trade itself. In general, we expect that a trade moves the price against the buyer or seller. That is, the price is pushed upwards when buying and downwards when selling. The market impact portion of slippage can be substantial when the ratio of the trade size to the average trade volume is high and is often modeled as an increasing function of this quantity.

Furthermore, based on the liquidity of different securities, their trading costs may be significantly different. Out of two securities with similar expected return and risk profiles, one with higher liquidity is more likely to have higher post-transaction cost returns. Therefore, a portfolio construction framework that ignores transaction costs will lead to suboptimal portfolios. For this purpose portfolio managers use forecasting models to predict the resulting transaction costs as they construct their portfolio. Several well-known models have been proposed in the literature, see, for example, [Hasbrouck \(1991\)](#), [Lillo, Farmer, and Mantegna \(2003\)](#), and [Almgren, Thum, Hauptmann, and Li \(2005\)](#).

To illustrate we focus our attention to the model by [Almgren et al. \(2005\)](#) whose main feature is that it explicitly and separately estimates the permanent ( $I_t^{\text{perm}}$ ) and temporary ( $I_t^{\text{temp}}$ ) market impacts for each order of  $x_i$  shares of stock  $i$

$$I_t^{\text{perm}}(x_i) = \gamma \cdot T \cdot \sigma_i \cdot \text{sign}(x_i) \cdot \left| \frac{x_i}{V_i \cdot T} \right|^\alpha \cdot \left( \frac{\Theta_i}{V_i} \right) + \mathcal{E}_i^{\text{perm}}$$

$$I_t^{\text{temp}}(x_i) = \eta \cdot \sigma_i \cdot \text{sign}(x_i) \cdot \left| \frac{x_i}{V_i \cdot T} \right|^\beta + \mathcal{E}_i^{\text{temp}}$$

<sup>1</sup> See, for example, [Fabozzi et al. \(2007\)](#) for a discussion of numerical optimization software for solving portfolio optimization problems.

<sup>2</sup> See also the interesting contribution to this debate in [Ceria \(2012\)](#).

<sup>3</sup> Capital gains taxes may make a strategy that is profitable on a pre-tax basis into a losing one on an after-tax basis. Therefore, it may be important for the portfolio manager to factor tax consequences into the investment decision. What makes this problem more complicated is the fact that, different investors are in different tax brackets and therefore have different effective tax rates. The after-tax allocation problem is a large topic by itself and is beyond the scope of this article. A good starting point for this topic is Don Mulvihill's articles "Investing for Real After-Tax Results," "Asset Allocation And Location," and "Equity Portfolio Structure" in [Litterman \(2003\)](#) Additional analysis on the impact of taxes can be found in [Stein and Garland \(2008\)](#), [Stein \(2001\)](#), and [Brandes, Domowitz, and Serbin \(2012\)](#).

where  $V_i$  is the stock's average daily volume,  $\sigma_i$  is the one-day standard deviation of the stock's return,  $\Theta_i$  is the number of outstanding shares of stock  $i$ ,  $T$  is the fraction of the day over which the trade is executed, and  $\varepsilon_i^{\text{perm}}$  and  $\varepsilon_i^{\text{temp}}$  are unexplained residual terms. The dimensionless term  $\Theta_i/V_i$  in the formulation of the permanent impact costs measure the fraction of the company's value traded each day and, as such, is a measure of relative liquidity of the stock.

Using a large set of trades, the cross-sectional model parameters  $\alpha, \beta, \gamma, \delta, \eta$  can be estimated, giving the following qualitative results.<sup>4</sup> First, permanent impact cost is linear ( $\hat{\alpha} \cong 1$ ) in trade size. Second,  $\hat{\beta} \cong 1/2$  meaning that the temporary impact cost is roughly proportional to the square root of the fraction of volume represented by one's own trading during the period of execution.<sup>5</sup> Hence, for a given rate of trading, a less volatile stock with large average daily volume has the lowest temporary impact costs.

Using this market impact model of the dollar transaction costs of a trade of  $x_i$  shares in security  $i$  become

$$TC_i(x) = |x_i| \cdot \left[ \frac{1}{2} I^{\text{perm}}(x_i) + I^{\text{temp}}(x_i) \right]$$

We observe that the resulting transaction cost function is of the form

$$TC_i(x_i) = a_i \cdot |x_i|^{3/2} + b_i \cdot x_i^2$$

With Almgren et al.'s choice of the exponent for the temporary impact term, one gets a slightly different functional form:  $TC_i(x_i) = a_i \cdot |x_i|^{8/5} + b_i \cdot x_i^2$ . For a basket of trades  $(x_1, \dots, x_n)$  the total dollar transaction costs therefore is  $TC(x) = \sum_i TC_i(x_i)$ .

We are now ready to write out an extension of the MVO that incorporates market impact costs. Suppose that the current portfolio holdings in dollars are  $\omega_0$  and we want to determine the new holdings  $\omega$ , given a vector of expected returns  $\alpha$ , a covariance matrix of returns  $\Sigma$ , and a risk aversion coefficient  $\lambda$ . The mean-variance problem with market impact costs then takes the form

$$\begin{aligned} \max_{\omega \in \Omega} \quad & \mu^\top \omega - \lambda \cdot \omega^\top \Sigma \omega - \gamma \cdot TC(x) \\ \text{s.t.} \quad & \omega^\top e + TC(x) \leq \omega_0^\top e \\ & w - x \circ p = w_0 \end{aligned}$$

where  $e = (1, 1, \dots, 1)^\top$ ,  $x$  are the trades expressed in terms of shares traded,  $p$  are the security prices and  $\gamma$  is a transaction cost aversion parameter. The expression  $x \circ p$  is the component-wise product of the trade and price vectors, representing the dollar value of the trades. The objective function combines the standard mean-variance utility with a market impact cost function. The budget constraint states that the market impact costs have to be financed from existing holdings.

Although this is a specific example, it is in general true that many realistic models for transaction costs involve nonlinear (and non-quadratic) functions. While the portfolio allocation problem without transaction costs is a quadratic program (QP) – with transaction costs it is a more general *nonlinear program* (NLP) and thus can be more time-consuming to solve, especially for a large number of assets.

There are two main approaches to handle this. One can proceed as in, for example, Ceria, Takriti, Tierens, and Sofianos (2008) and use a specialized solver (second-order cone programming (SOCP)

solvers) to directly solve these type of problems, or one can solve a QP relaxation of the problem. We note that in principle there is no loss in numerical accuracy by solving a problem where the market impact function has been approximated by a simple linear/quadratic function. This is due to the fact that estimated market impact models have *large* estimation errors. A benefit of the approximation is that it can be solved very quickly with standard QP software.

In the direct solution approach, we can incorporate 3/2-power transaction cost functions (coming from square root market impact cost functions) directly into a second-order cone programming solver. This fact is based on the simple observation that the following two optimization problems are equivalent:

$$\min_{x_1, \dots, x_N} \sum_{i=1}^N a_i x_i^{3/2}$$

and

$$\begin{aligned} \min_{x_1, \dots, x_N, y_1, \dots, y_N} \quad & \sum_{i=1}^N a_i y_i \\ \text{s.t.} \quad & x_i^{3/2} \leq y_i, \dots, x_N^{3/2} \leq y_N. \end{aligned}$$

Next, we note that the constraints are equivalent to the rotated quadratic cone inequalities  $x_i^2 \leq y_i \cdot z_i$  and  $z_i^2 \leq x_i$ . This last equivalence is well known and is noted, for example, in Exercise 9.2 of Cornuejols and Tütüncü (2007). It is also possible to represent the market impact functions of the form  $TC_i(x_i) = a_i \cdot |x_i|^{8/5} + b_i \cdot x_i^2$  in conic form after applying similar but more complex variable substitutions.

Further discussion of the transaction costs in the portfolio optimization can be found in, for example, Borkovec, Domowitz, Kiernan, and Serbin (2010), Fabozzi, Focardi, and Kolm (2010), Chen, Fabozzi, and Huang (2010), Brandes et al. (2012), and references therein.

### 3.2. Constraints in portfolio construction

Starting with the first papers of Markowitz on the subject, most MVO formulations featured a variety of constraints such as the long-only investing constraint. Indeed, one of the appeals of MVO in quantitative portfolio construction is the flexibility this approach offers for incorporating multiple constraints reflecting client guidelines, regulatory restrictions, as well as the discretionary views of the portfolio manager. The algorithms and software used for MVO are well equipped to deal with a variety of constraint types in a generic and robust manner. This flexibility allows the portfolio manager to customize the portfolios according to different client requirements and risk appetites as well as the varying regulatory restrictions of the different jurisdictions where the products are offered without needing to alter the alpha and risk signals used for building the portfolios.

However, the presence of constraints impacts the portfolio construction process as well as the characteristics of the constructed portfolios and their performance in material and often unpredictable ways. Using constraints without understanding and quantifying this impact can lead to poor portfolios that do not reflect the views and expertise of the portfolio manager. For this reason, a number of studies in recent years focused on building diagnostic measures and tools to evaluate the impact of constraints on portfolios and their performance.

Clearly, adding constraints to an MVO problem can never improve the ex ante optimization results. However, in practice the inclusion of constraints in the mean-variance optimization problem can lead to better out-of-sample performance, compared

<sup>4</sup> The authors use a proprietary data set obtained from Citigroup's equity trading desk in which a trade's direction – buyer or seller initiated – is known. Note that for most public data sets, trade direction is not available and has to be estimated by a classification algorithm. Classification errors in algorithms introduce a bias that typically produces an overestimate of the true trading cost (Ellis, Michaely & O'Hara, 2000; Lee & Ready, 1991).

<sup>5</sup> Almgren et al. actually find that  $\hat{\beta} \cong 3/5$ . However, many practitioners prefer 1/2 partly because it leads to a more tractable formulation.



to portfolios constructed without these constraints. For example, practitioners often use long-only constraints or upper and lower bounds for each security to avoid overconcentration in a few assets. Gupta and Eichhorn (1998) suggest that constraining portfolio weights may also assist in containing portfolio volatility, increase realized portfolio performance, and decrease downside risk and shortfall probability. These observations are related to the *model insurance* property of certain constraints—we discuss this in more detail below.

We start the section with an informal taxonomy of constraints used in portfolio construction, especially with respect to the source and motivation of these constraints. We then review the concept of the *transfer coefficient* as well as the quantification of the cost of constraints using the well-known notion of *shadow costs* from the optimization literature. We also discuss the interesting contributions to the literature that study the interactions between constraints and the misalignment between alpha and risk models.

### 3.2.1. Regulatory constraints

Some of the constraints used in portfolio construction reflect the restrictions imposed on the portfolio manager by the market regulators. As such, they are inflexible and must be respected at all times, even if they limit the portfolio manager's ability to add value to the portfolio through security selection and positioning. For example, mutual fund managers in the US are subject to the restrictions of the *Investment Company Act of 1940*, regulated and enforced by the Securities and Exchange Commission. Similarly, in many European jurisdictions, portfolio managers are bound by the *Undertakings for Collective Investments in Transferable Securities* (UCITS) directive. In addition to these rules that intend to regulate the actions of portfolio managers, regulators occasionally issue temporary limitations for risk management purposes. Examples of such constraints include the short-selling limitations imposed in various jurisdictions.<sup>6</sup>

### 3.2.2. Guideline constraints

When a portfolio manager is managing money for a client, the client may specify certain restrictions on the portfolio's positions or trades. Such restrictions are usually termed *guideline restrictions*, and are carefully specified in an investment management agreement (IMA). For example, the client may ask the manager not to invest in certain industries such as tobacco, or may impose a limit on the number of holdings in the portfolio. Like the regulatory restrictions, guideline restrictions are considered to be inflexible and cannot be violated without the client's consent.

### 3.2.3. Discretionary exposure constraints

Frequently, the constraints used in portfolio construction are discretionary and may originate from the portfolio manager's desire to limit exposure to certain securities, or groups of securities such as those in a particular industry, sector, or country. Similarly, the manager may want to limit exposure to securities with certain characteristics such as small or large market capitalization, high or low leverage, or common quantitative factor exposures such as value and momentum.

A well constructed and calibrated risk model will accurately price the risk in taking such exposures and the MVO process will yield the optimal risk-return tradeoffs for the portfolio. So, why does the portfolio manager need to use exposure constraints that may inhibit the ability of the optimization process to generate ideal tradeoffs? The answer to this question is related to the notion of *model insurance*. The portfolio manager recognizes that any

model of security returns and risks are subject to estimation errors and may underestimate the risks a portfolio may be exposed to. Underestimation of risks can lead to extreme bets in an optimized portfolio and the portfolio manager may therefore instead choose to limit the size of the bets explicitly using constraints.

### 3.2.4. Trading constraints

As we discussed in Section 3.1, trades that constitute a large fraction of the total trading volume of a security tend to move the price of the security in an unfavorable direction and incur market impact costs. In addition to, or instead of, incorporating these costs into the portfolio selection methodology, a portfolio manager may also choose to explicitly limit the participation rate in trading certain securities. These constraints typically have the following form: "do not trade more than  $x\%$  of the average daily volume of security  $y$ ".

### 3.2.5. Risk management constraints

Portfolio managers often limit the predicted volatility of the portfolio, measured either in an absolute sense or with respect to a benchmark. In addition, risk contributions of individual securities or groups of securities may be restricted. Such constraints are useful for risk budgeting purposes, for example, for allocating risk across different geographic regions or different risk factors in an equity portfolio.

### 3.2.6. Transfer coefficient

Our ability to understand and quantify the impact of constraints on portfolio performance and characteristics has improved dramatically since the turn of the century. One of the pioneering works in this field is that of Clarke, De Silva, and Thorley (2002). This work introduced the concept of the *transfer coefficient* which is defined to be the cross-sectional correlation coefficient between the risk-adjusted active weights in an optimized portfolio and the risk-adjusted forecasted active returns ("alphas") for the corresponding securities.

Clarke et al. (2002) motivate the definition of the transfer coefficient using the observation that unconstrained optimal risk adjusted active weight allocations are proportional to the alphas when active returns are assumed to be uncorrelated. This corresponds to a transfer coefficient of 1 for the portfolio and provides a useful base case for comparison with the more general settings. For example, when the portfolio optimization problem includes constraints, risk-adjusted active weights are no longer proportional to the alphas and the transfer coefficient depends on how well these weights match their alphas.

As we argued above, most portfolio optimization problems in practice have multiple constraints including regulatory, guideline, and discretionary limits on positions or trades. One can associate a cost with a constraint by measuring this distortion, expressed as the transfer coefficient of the resulting portfolio. In their study, Clarke et al. (2002) compute the transfer coefficient in various scenarios involving the long-only constraint, a market-cap neutrality constraint, and a turnover constraint. Not surprisingly, their analysis shows that the long-only constraint is the costliest of these constraint types.

### 3.2.7. Shadow cost decomposition

One important limitation of the approach in Clarke et al. (2002) is that when a portfolio is constructed using multiple constraints, it only provides an aggregate measure of the impact of these constraints and does not offer a scalable strategy to decompose the effects of individual constraints. Therefore, it becomes harder to assess the costliness of each constraint in realistic settings using transfer coefficients. In recognition of this limitation, several studies attempt to decompose the impact of individual constraints,

<sup>6</sup> See, for example, Lee and Stefek (2008) who discuss the short-selling bans imposed by regulators in Spain and Italy.

often using the notion of *shadow costs* (or, *Lagrange multipliers*) from the optimization literature. This approach is based on the first-order optimality conditions for constrained optimization problems. Given its heavy reliance on shadow costs, this methodology is sometimes called the *shadow cost decomposition*. Tütüncü (2012) demonstrates the applications of shadow cost decomposition in utility analysis, active-weight decomposition, return factor analysis, and performance attribution.

Previous work using this representation includes Grinold (2005) where the decomposition allows the allocation of *opportunity loss* to different costs and constraints in the portfolio construction problem. In this work, the opportunity loss is measured in terms of Grinold's objective function which is a simple risk-adjusted portfolio return function. We note that this study focuses more on the impact of trading costs than the constraints. Scherer and Xu (2007) focuses on the impact of constraints on investor utility. Scherer and Xu (2007). They make the observation that while the impact of constraints on individual security weights can be severe, this impact may be less pronounced on the utility. They argue that the constraint impact is best measured in the utility dimension and they provide this measurement using the decomposition above.

Bender, Lee, and Stefek (2009) find additional insights by splitting the portfolio weight distortions caused by a constraint into two parts: one that is aligned with the unconstrained optimal portfolio and one that is orthogonal to it.

Stubbs and Vandenbussche (2010) also analyze the shadow cost decomposition for measuring the impact of constraints. They extend the methodology by developing an ex-post constraint attribution on optimized portfolios by attributing realized returns to objective terms and constraints, as opposed to sectors or factors that are used in classical performance attribution. They achieve this essentially by tracking the performance of the “distortion portfolios” associated with constraints in the shadow cost decomposition approach.

The return-based performance analysis developed by Stubbs and Vandenbussche (2010) complements the *ex ante* analysis described in the other references. One of the questions this constraint attribution method can answer is related to the *model insurance* function of constraints we discussed above. As we argued, a portfolio manager may explicitly limit exposures to certain securities or groups of securities reflecting the lack of conviction in a model that may produce optimal weights beyond such limits. On an *ex ante* basis, such constraints can never improve the utility of the optimized portfolio and will often decrease it. The *ex post* constraint attribution analysis may actually reveal that an insurance constraint improves performance by preventing the portfolio from taking bets that turn out to be harmful to the portfolio performance. Or it may confirm that the constraint is harmful to the portfolio performance, either because the “insurance premium” is too high or that the constraint does not actually provide sufficient protection. As with any insurance strategy, such an evaluation must be made on an on-going basis and use observations from a relatively long history.

### 3.2.8. Misalignment from constraints

The concept of misaligned alpha and risk models has attracted attention in the recent literature; see, for example, Lee and Stefek (2008) and Ceria, Saxena, and Stubbs (2012). These studies demonstrate that when the alpha model contains factors that are not priced (or, spanned) by the risk model, then the optimization process will exploit this fact and generate portfolios that provide excess returns without incurring risks, as measured by the risk model. The presence of such arbitrage opportunities makes the optimization process unstable and the resulting portfolios have the property that the risk of the portfolio is significantly under-predicted. This under-prediction is caused by the component of the

alpha model that is orthogonal to the risk model. By ensuring that the alpha and risk models are aligned correctly, or that all alpha model factors are captured in the risk model as well, some of the issues arising from misalignment can be avoided.

In a recent contribution to this debate, Saxena, Martin, and Stubbs (2012) analyze the misalignments caused by constraints. They observe that even when the alpha model and the risk model are aligned correctly and the alpha model has no component that is orthogonal to the risk model, if binding constraints are used in portfolio construction, resulting optimal portfolios can have significantly under-predicted levels of active risk. To be more precise, the realized active risks of these portfolios can be 20–30% above the levels of risk predicted by their risk model.

Saxena et al. (2012) consider the case of active weight bound constraints and attribute the risk estimation bias caused in this scenario to the presence of systematic risk exposures not captured by the risk model in assets for which the active weight constraints are binding. In other words, the combination of the active weight constraint and the use of the MVO methodology reveal a systematic risk factor that is missing from the risk model. The authors offer the *alpha alignment factor* methodology to address this serious issue. The methodology works by creating disincentives for exposures in the hidden risk factor and thus delivers portfolios with unbiased risk estimates.

### 3.3. Mitigating the impact of estimation errors in MVO

In the classical MVO framework an investor is required to provide estimates of the expected returns and covariances of all the securities in the investment universe considered. This is a formidable task given the number of securities available today. Portfolio managers are unlikely to have a detailed understanding of all the securities, companies, industries, and sectors that they have at their disposal. Typically, most of them have a specific area of expertise that they focus on in order to achieve superior returns. This is likely to be one of the major reasons why the MVO framework has not been adapted by more practitioners. It is simply unrealistic to expect the portfolio manager to produce reasonable estimates (besides the additional problems of estimation error) of the inputs required in classical portfolio theory.

In addition, after the estimation of these quantities, the portfolio optimization problem is often solved as a *deterministic* problem – thereby completely ignoring the uncertainty in the inputs. Using point estimates of the expected returns and the covariance matrix of returns, and treating them as error-free in portfolio allocation is not prudent. If, say, the portfolio manager has more confidence in some of the estimates than others, it may make sense to treat these inputs differently while building portfolios. With these concerns in mind, some of the recent approaches to portfolio optimization consider the incorporation of the uncertainty of expected returns and risk into the optimization process to create a more realistic model.

In a portfolio optimization context, securities with large expected returns and low standard deviations will be overweighted in comparison to their benchmark weights. Conversely, securities with low expected returns and high standard deviations will be underweighted. Consider the scenario where the expected returns and covariances are estimated and the estimation errors for these quantities are randomly distributed with mean zero. This means that some quantities are over-estimated and others are underestimated. The optimization process will allocate higher weights to securities with over-estimated expected returns and under-estimated risks, and symmetrically, lower weights to securities with under-estimated expected returns and over-estimated risks. The larger the estimation error, the larger the impact will be on optimized weights. For this reason, some authors cynically refer to optimizers as “error maximizers.” While the impact of small

estimation errors on portfolio weights can indeed be significant, this often happens when some assets are close substitutes for another (Kritzman, 2006). As a result, portfolio return distribution does not change dramatically even when the optimal weights are changing. Using this observation, Kritzman argues that mean–variance optimizers are often robust to estimation errors when their sensitivity is measured in the portfolio return space as opposed to portfolio weights. While the debate around the “error maximization” property continues, the research around understanding and mitigating the impact of estimation errors has reached a mature state.

Uncertainty from estimation error in expected returns tends to have more influence than in the covariance matrix in a MVO (Best & Grauer, 1991a, 1991b). The relative importance depends on the investor's risk aversion, but as a general rule of thumb, errors in the expected returns are more important than errors in the covariance matrix, and errors in the variances are about twice as important as errors in the covariances (Chopra & Ziemba, 1993; Kallberg & Ziemba, 1984). As the risk tolerance increases, the relative impact of estimation errors in the expected returns becomes even more important. Conversely, as the risk tolerance decreases, the impact of errors in expected returns relative to errors in the covariance matrix becomes smaller. As a rule of thumb, the first priority should be on providing good estimates for the expected returns, followed by the variances and correlations. In this section, we discuss some of the most common techniques for mitigating estimation errors.

### 3.3.1. Constraints on portfolio weights

As we argued in Section 4.2, portfolio managers often impose limits on the portfolio weights of securities or groups of securities to avoid extreme weights that may result from model inaccuracies. Jagannathan and Ma (2003) provide a theoretical justification for such practices. They show that the no short-selling constraints are equivalent to reducing the estimated security covariances, whereas upper bounds are equivalent to increasing the corresponding covariances. For example, stocks that have high covariance with other stocks tend to receive negative portfolio weights. Therefore, when their covariance is decreased (which is equivalent to the effect of imposing no short-selling constraints), these negative weights diminish in magnitude. Similarly, stocks that have low covariances with other stocks tend to get over-weighted. Hence, by increasing the corresponding covariances the impact of these over-weighted stocks decrease. However, extreme care needs to be taken when imposing constraints for robustness and stability purposes. For example, if the constraints used are too “tight,” the portfolio allocation will be completely determined by the constraints instead of the forecasted expected returns and their covariances.

### 3.3.2. Diversification measures

Instead of providing ad hoc upper and lower bounds on each security, Bouchard, Potters, and Aguilar (1997) suggest using “diversification indicators” that measure the concentration of the portfolio. These diversification indicators can be used as constraints in the portfolio construction phase to limit the concentration to individual securities. The authors demonstrate that these indicators are related to the information content of the portfolio in the sense of information theory. In particular, a concentrated portfolio corresponds to a large information content (as we would only choose a very concentrated allocation if our information about future price fluctuations is “perfect”), whereas an equally weighted portfolio would indicate low information content (as we would not put “all the eggs in one basket” if our information about future price fluctuations is poor).

Constraints on risk contributions of individual securities or groups of securities we discussed in Section 3.2 are also examples of approaches that aim to mitigate the impact of estimation errors by introducing constraints on diversification metrics.

### 3.3.3. Bayesian techniques and the Black–Litterman model

A great number of general Bayesian and shrinkage approaches have been used to estimate the inputs to MVO. For the expected returns see, for example, Jobson and Korkie (1981), Frost and Savarino (1986), and Jorion (1991, 1986). For the covariance matrix see, for example, Ledoit and Wolf (2003, 2004).<sup>7</sup> The basic idea underlying these types of estimators is the *bias-variance tradeoff*,<sup>8</sup> where by sacrificing some bias one can obtain a more efficient estimator that is less sensitive to changes in the data.

In this section we focus on the Black–Litterman model that is a “market based” shrinkage approach where the estimate of expected returns is calculated as a weighted average of the market equilibrium (e.g., the CAPM equilibrium) and the investor's views. The weights depend upon (1) the volatility of each asset and its correlations with the other assets and (2) the degree of confidence in each forecast.

Let us mention at this point that the ability to incorporate exogenous insight, such as a portfolio manager's judgment, into formal models is important in practice. Such insight might be a valuable input in the model. The Bayesian framework allows forecasting systems to use such external information sources and subjective interventions (i.e., modification of the model due to judgment) in addition to traditional information sources such as market and proprietary data. Because portfolio managers might not be willing to give up control to a “black box,” the incorporation of exogenous insights into formal models through Bayesian techniques is one way of giving the portfolio manager better control of a quantitative framework. Forecasts are represented through probability distributions that can be modified or adjusted to incorporate other sources of information deemed relevant. The only restriction is that such additional information (i.e., the investor's “views”) be combined with the existing model through the laws of probability. In effect, incorporating Bayesian views into a model allows to “rationalize” subjectivity within a formal quantitative framework. “[T]he rational investor is a Bayesian,” as Markowitz noted.<sup>9</sup>

Many trading strategies cannot easily be turned into forecasts of expected returns and covariances. In particular, trading strategies may not produce views on *absolute* return, but rather just provide *relative* rankings of securities that are predicted to outperform/underperform other securities. For example, let us consider two stocks, A and B. An *absolute view* is of the form “the one-month expected return on A and B are 1.2% and 1.7% with a standard deviation of 5% and 5.5%, respectively.” In contrast, a *relative view* may be of the form “B will outperform A by half a percent over the next month” or simply “B will outperform A over the next month”. Clearly, it is not an easy task to translate any of these relative views into the inputs required in the modern portfolio theoretical framework.

An assumption underlying the Black–Litterman model is that the expected return of a security should be consistent with market equilibrium *unless* the investor has a specific view of the security. In other words, an unconstrained investor who does not have any views on the market should hold the market.

<sup>7</sup> While less popular in practice, robust statistics approaches are also used to estimate the inputs to MVO (see, for example Trojani and Vanini (2002) and DeMiguel and Nogales (2009)). Based on space constraints we have omitted a discussion of these techniques in this paper.

<sup>8</sup> It can be shown that the mean squared error of an estimator is equal to its variance plus its squared bias. The minimum of the mean squared error may not be obtained when the bias is zero.

<sup>9</sup> See page 57 in Markowitz (1987).

We assume that security returns a multivariate normally distributed with mean  $\mu$  and covariance matrix, i.e.  $r \sim N(\mu, \Sigma)$ . However,  $\mu$  is itself a multivariate normally distributed random vector  $\mu \sim N(\pi, \Sigma_\pi)$  where  $\pi$  is a vector of market equilibrium returns (more about this below). These assumptions encapsulate the idea that the expected returns of the securities deviate from the perceived market equilibrium. Naturally, both the expected return vector and the vector of market equilibrium returns are not observable and have to be estimated. We now illustrate how to use the Black–Litterman model in three steps.<sup>10</sup>

**Step 1: Investor Views.** An investor may have views on some or all of the securities. The investor expresses these views as  $P\mu \sim N(q, \Omega)$ . Here the matrix  $P \in \mathbb{R}^{k \times n}$  describes the investor's views, the vector  $q \in \mathbb{R}^k$  are the expected returns of the investor's views ("alphas"), and  $\Omega \in \mathbb{R}^{k \times k}$  is the covariance matrix of the views (the "confidence"). The matrix  $P$  is sometimes referred to as the "picking matrix" as it "picks out" the securities that the investor has views about.

**Step 2: Market Equilibrium.** We need to estimate market equilibrium,  $\pi$ . The standard approach in the Black–Litterman model is to use the Capital Asset Pricing Model (CAPM), that is

$$\pi_i \equiv E(r_i) - r_f = \beta_i(E(r_M) - r_f)$$

$$\beta_i = \frac{\text{COV}(r_i, r_M)}{\sigma_M^2}$$

where  $E(r_i)$  is the expected return on security  $i$ ,  $E(r_M)$  is the expected return on the market portfolio,  $r_f$  is the risk-free rate,  $\beta_i$  is the beta of security  $i$ , and  $\sigma_M^2$  is the variance of the market portfolio. Denoting by  $\omega_b = (\omega_{b1}, \dots, \omega_{bn})$  the market capitalization or benchmark weights, CAPM can be written as

$$\pi = \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_n \end{bmatrix} = \delta \Sigma \omega$$

$$\delta = \frac{E(r_M) - r_f}{\sigma_M^2}$$

where  $\delta$  is the market price of risk. In this context  $\pi$  is referred to the vector "market implied expected returns" or "equilibrium implied expected returns."

**Step 3: The Black–Litterman Expected Returns.** The Black–Litterman expected returns are given by the formula

$$\hat{\mu}_{BL} = [(\tau \Sigma)^{-1} + P^\top \Omega^{-1} P]^{-1} [(\tau \Sigma)^{-1} \pi + P^\top \Omega^{-1} q]$$

with covariance matrix ("uncertainty")

$$M \equiv \text{Var}(\mu_{BL}) = [(\tau \Sigma)^{-1} + P^\top \Omega^{-1} P]^{-1}$$

This is the Black–Litterman model for the market equilibrium combined with the investor's views.

Some remarks are in order to provide a better intuitive understanding of the Black–Litterman formula. First, we see that if the investor has no views or the confidence in the views is zero (i.e.  $q = 0$  or  $\Omega = 0$ ), then the Black–Litterman returns are equal to the equilibrium returns, i.e.  $\hat{\mu}_{BL} = \pi$ . Consequently, with no views the investor will end up holding the market portfolio. Second, it is not hard to see that the Black–Litterman expected returns are a "confidence" weighted linear combination of market equilibrium and the investor's views with the two weighting matrices

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

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

and

$$\omega_\pi + \omega_q = I$$

Specifically,  $(\tau \Sigma)^{-1}$  and  $P^\top \Omega^{-1} P$  represent the confidence we have in the estimates of the market equilibrium and views, respectively. Therefore, if we have low confidence in the views, the resulting expected returns will be close to the ones implied by market equilibrium. Conversely, with higher confidence in the views, the resulting expected returns will deviate from the market equilibrium implied expected returns. We say that we "tilt" away from market equilibrium.

It is straightforward to show that the Black–Litterman expected returns can also be written in the form

$$\hat{\mu}_{BL} = \pi + \tau \Sigma P^\top [P \tau \Sigma P^\top + \Omega]^{-1} [q - P \pi]$$

From this expression we see that the "tilt" away from the equilibrium is given by a vector proportional to  $\Sigma P^\top [P \tau \Sigma P^\top + \Omega]^{-1} [q - P \pi]$ .

Probably one the most important features of the Black–Litterman model is that it "adjusts" the entire market equilibrium implied expected return vector with the investor's views. Because security returns are correlated, views on just a few assets will imply changes to the expected returns on all assets. Mathematically speaking, this follows from the fact that although the vector  $q$  can have dimension  $K \ll N$ ,  $P^\top \Omega^{-1} P$  is an  $N \times K$  matrix that "propagates" the  $K$  views into  $N$  components,  $P^\top \Omega^{-1} P q$ . 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 forecasts will be interpreted as an arbitrage opportunity by a mean–variance optimizer resulting in portfolios concentrated in just a few assets ("corner solutions").

There are a number of extensions to the Black–Litterman model. First, we note that it may appear that the equilibrium-based prior used in Black–Litterman restricts the potential applications of the technique to global and diversified funds. This is not true. The posterior formula (the Black–Litterman expected return vector and covariance matrix) can be applied to *any* normal distribution (as the prior), not just market equilibrium. As pointed out in [Meucci \(2010\)](#) active management was among the first applications of Black–Litterman, where the prior expectation was assumed to be null.

[Satchel and Scowcroft \(2000\)](#) propose a model where an investor's view on global volatility is incorporated in the prior views by assuming that  $\tau \ll 1$  is unknown and stochastic. [Qian and Gorman \(2001\)](#) describe a technique based on conditional distribution theory that allows an investor to incorporate views on any or all variances. [Meucci \(2008\)](#) propose a methodology called the "entropy pooling approach" to incorporating nonlinear views in a non-normal market. Here, views can be seen as statements that distort the prior distribution of security returns. The posterior distribution is obtained by combining the views and prior distribution in such a way that the entropy of the posterior distribution is minimized relative to the prior.

### 3.3.4. Robust optimization

The sensitivity of the output of optimization algorithms, including MVO, to changes in problem inputs is well-documented and is related to the often discontinuous and ill-behaved mapping between the inputs and outputs of optimization problems ([Bonnans & Shapiro, 2000](#)). In addition, constraints in

<sup>10</sup> For a detailed description and derivation see, for example, [Fabozzi et al. \(2007\)](#).



the problem formulation may amplify the sensitivity of the results to the changes in the inputs. Combined with the fact that inputs to MVO are subject to estimation errors, these general properties of optimization algorithms lead to concerns around the reliability of the decisions derived using portfolio optimization methods.

During the last 15 years, robust optimization methodology emerged as an answer to some of these concerns and refers to the modeling of optimization problems with data uncertainty in such a way that one finds a solution that is guaranteed to be satisfactory for most realizations of the uncertain parameters (Ben-Tal & Nemirovski, 1998; Ben-Tal & Nemirovski, 1999; El Ghaoui & Lebret, 1997; El Ghaoui, Oustry, & Lebret, 1998). *Uncertainty sets* that contain possible values of the uncertain parameters are used to describe the uncertainty in the problem and their size represents the level of the uncertainty and/or the desired level of robustness.

Once an uncertainty set is defined, the robust portfolio optimization problem is formulated using an adversarial perspective. In this formulation, once the investor makes a portfolio selection, it is assumed that an adversary will choose the expected returns and covariances from the uncertainty set so as to minimize the investor's utility. Another way to view the choice of the adversary is to think of it as the worst case realization of the input parameters (expected returns and covariances) from their uncertainty set. In this setting, the objective of the investor is to choose a portfolio that maximizes the worst-case utility (Goldfarb & Iyengar, 2003; Lobo & Boyd, 2000; Tütüncü & Koenig, 2004). In addition to robust counterparts of the MVO framework, there are robust formulations that are based on additional risk measures such as value-at-risk (El Ghaoui, Oks, & Oustry, 2003) and conditional value-at-risk (Garlappi, Uppal, & Wang, 2007; Zhu & Fukushima, 2009) also introduce methods to increase robustness by incorporating factor models.

Robust optimization formulations lead to challenging mathematical problems but in many cases modern optimization techniques such as second-order cone optimization or semidefinite optimization provide tools and software that make robust portfolio optimization problems computationally tractable. Due to the complexity of robust formulations, there also have been efforts to understand the properties of robust portfolios (Gregory, Darby-Dowman, & Mitra, 2011; Kim, Kim, Ahn, & Fabozzi, 2012; Kim, Kim, & Fabozzi, 2012, 2013, in press; Kim, Kim, Kim, & Fabozzi, 2014), additional details on robust optimization and its use in portfolio selection are provided in Fabozzi et al. (2007), Cornuejols and Tütüncü (2007), Fabozzi, Huang, and Zhou (2010), and Kim et al. (in press).

As with any attempt to improve mean–variance optimization, robust optimization is not without its critics. For example, Scherer (2007) argues that the robust optimization methods are essentially equivalent to using shrinkage estimators on the inputs and, at least in certain scenarios, lead to portfolios with inferior out-of-sample results.

### 3.3.5. Incorporating higher moments and tail-risk measures

The mean–variance framework is a special case of so-called expected utility maximization where investors are assumed to have quadratic utility or returns distributions are jointly normal. This may sometimes be limiting as many financial return distributions are not jointly normal, but exhibit fat tails and asymmetry that cannot be described by their mean–variances alone. In many instances, the tails of the return distribution significantly affect portfolio performance (see, for example, Jobst and Zenios (2001)). Harvey and Siddique (2000) show that skew in stock returns can be relevant in portfolio selection. In particular, if asset returns exhibit non-diversifiable co-skew, investors expect be rewarded for

it, resulting in increased expected returns. They illustrate that in the presence of positive skew, investors may be willing to accept a negative expected return.

Given the computational power available today it is possible to construct portfolios (at least portfolios of moderate size) that maximize expected utility under the empirical asset return distribution. In practice however, this approach is seldom used. Typically, practitioners rely upon mean–variance approximations of a chosen utility function. Levy and Markowitz (1979) compared the performance of portfolio allocation by maximizing expected power utility with that of the standard mean–variance optimization. They found that mean–variance approximations often perform well enough. Cremers, Kritzman, and Page (2003) and Cremers, Kritzman, and Page (2005) show empirically that the log and power utility functions are fairly insensitive to higher moments, and therefore, mean–variance optimization performs well for investors with log or power preferences. However, for discontinuous or S-shaped utility functions (for example, such as that of prospect theory) this result no longer holds true and mean–variance optimization shows significant loss in utility compared to an optimization of the full utility function.

One alternative to general expected utility maximization is to extend the mean–variance framework by directly incorporating portfolio skew and kurtosis. In fact, such extensions can be seen as approximations to general expected utility maximization, where one considers a Taylor series expansion of the utility function and drops the higher order terms from the expansion. The first attempt to extend the classical mean–variance optimization in this fashion was done by Jean (1971). Later, more general and rigorous treatments have been presented by several authors (see for example, Athayde and Flôres (2002, 2004) and Harvey, Liechty, Liechty, and Mueller (2010)).

Another alternative to general expected utility maximization is to maximize expected portfolio return subject to some tail-risk measure. We note that in this case the problem may not be equivalent to that of maximizing a utility function.

Probably the most well-known risk measure, besides the standard deviation, is the Value-at-Risk (VaR) first developed by J.P. Morgan and made available through the RiskMetrics™ software in October 1994 (see, Morgan (1996)). While VaR is popular as a risk measure, it has several undesirable mathematical characteristics such non-subadditivity and non-convexity. This led to the introduction of so-called coherent risk measures (see, Artzner, Delbaen, Eber, and Heath (1999)). Today, one of the most popular coherent risk measures used for portfolio construction is that of Conditional Value-at-Risk (CVaR). Rockafeller and Uryasev (2000) showed that these types of problems can be solved efficiently as linear programs (LPs).

## 4. New directions in portfolio optimization

In this section we selectively highlight some of the new trends and developments in MVO and its related areas. Due to space constraints, we cannot survey all new trends in this area. While admittedly our choice is subjective, it is based on what we believe are as some of the important developments in this area for the use of MVO and its extensions in practice.

First, we discuss the recent focus on diversification methods and provide a summary of the developments related to risk-parity portfolios. Second, we provide a formalization of the problem of mixing different sources of alpha and address some of the challenges that arise from these formulations. Finally, we outline some of the recent literature on the usage of multi-period portfolio optimization in practice.

#### 4.1. Diversification methods

One of the most important goals of quantitative portfolio management is diversification across sources of returns and risks in a portfolio. A simple way to achieve diversification is to allocate capital equally across a group of investment alternatives. Recent studies claim this simple approach can be effective and reach the maybe surprising conclusion that a variety of approaches aiming to achieve *optimal diversification* do not consistently perform better than the equally-weighted allocation according to multiple criteria in out-of-sample tests (DeMiguel et al., 2009). Besides its simplicity, an important “advantage” of the equally-weighted strategy is that it does not make use of return and risk models, and therefore, is not subject to the estimation errors in such models.

Other alternatives to achieve diversified portfolios include methods that make use of a risk model but no return model. One well-known example of these so-called *risk-based* allocation approaches is the global minimum risk portfolio, or the *global minimum variance portfolio* where risk is measured using the variance of the portfolio return. A related risk-based method is the *risk parity* approach (Asness, Frazzini, & Pedersen, 2012). The risk parity approach in portfolio construction aims to build portfolios where the overall portfolio risk is diversified by allocating the risk equally across the different investment strategies and/or securities. In this approach, the portfolio risk and the risk contributions are calculated from the variance and covariance estimates of their future returns.

##### 4.1.1. Measuring risk contributions

One of the important concepts in portfolio management is quantifying the risk of individual components (such as that of strategies and/or securities) to the total portfolio risk. The distribution of risk contributions from different portfolio components can then be used to measure the level of diversification within the portfolio.

There are different ways to define the risk contribution of an individual position to a portfolio. For example, one can consider the difference between the risk measures of the existing portfolio and the portfolio that is obtained by removing one of the positions and label that difference as the risk contribution of that position. To be more precise, define

$$\omega_{-i} = [\omega_1, \dots, \omega_{i-1}, 0, \omega_{i+1}, \dots, \omega_n]^T$$

which represents the portfolio without position  $i$ . Then, the risk contribution of position  $i$  can be defined as

$$\sigma_i(\omega) = \sigma(\omega) - \sigma(\omega_{-i})$$

This is the approach outlined for computing the *marginal risk* from a portfolio component in CreditMetrics (J. P. Morgan & Co., 1997). One problem with this definition of the risk contribution is that sum of the contributions of all the portfolio positions is in general not equal to the total portfolio risk, thus making the interpretation of the risk contributions unintuitive.

An alternative definition of risk contribution often used in practice is constructed as follows. First, define the marginal risk contribution (MRC) of asset  $i$  as the partial derivative of the function  $\sigma(\omega)$  with respect to  $\omega_i$ . This is the rate of change in the portfolio risk as the weight of asset  $i$  increases

$$\text{MRC}_i(\omega) = \frac{\partial \sigma(\omega)}{\partial \omega_i} = \frac{(\Sigma \omega)_i}{\sigma(\omega)}$$

Above,  $(\Sigma \omega)_i = \sum_{j=1}^n \sigma_{ij} \omega_j$  is the  $i$ th component of the vector  $(\Sigma \omega)$ . Then, the *risk contribution* (RC) of asset  $i$  is defined as the product of its weight and MRC

$$\text{RC}_i(\omega) = \omega_i \cdot \text{MRC}_i(\omega) = \frac{\omega_i \cdot (\Sigma \omega)_i}{\sigma(\omega)}$$

An important property of this definition is that the sum of the risk contributions for all assets in the portfolio is the total risk of the portfolio, that is

$$\sum_{i=1}^n \text{RC}_i(\omega) = \sum_{i=1}^n \frac{\omega_i \cdot (\Sigma \omega)_i}{\sigma(\omega)} = \frac{\omega^\top \Sigma \omega}{\sigma(\omega)} = \sigma(\omega)$$

We also define the *relative risk contribution* (RRC) of an asset to be the ratio of its risk contribution to the total portfolio risk, that is

$$\text{RRC}_i(\omega) = \frac{\text{RC}_i(\omega)}{\sigma(\omega)} = \frac{\omega_i \cdot (\Sigma \omega)_i}{\sigma^2(\omega)} = \frac{\omega_i \cdot (\Sigma \omega)_i}{\omega^\top \Sigma \omega}$$

##### 4.1.2. Risk parity

Given an  $n$ -dimensional covariance matrix  $\Sigma$ , a portfolio  $\omega = [\omega_1, \dots, \omega_n]^\top$  is called a *risk parity portfolio* with respect to  $\Sigma$  if it satisfies the following condition

$$\text{RC}_i(\omega) = \frac{\sigma(\omega)}{n}, \quad \text{for all } i = 1, \dots, n$$

Note that  $\omega$  is a risk parity portfolio if and only if

$$\text{RRC}_i(\omega) = \frac{1}{n}, \quad \text{for all } i = 1, \dots, n$$

In view of the last equation, risk parity portfolios can be compared to equal weighted portfolios where  $\omega_i = \frac{1}{n}$ , for all  $i = 1, \dots, n$ . Instead of allocating capital evenly across all the assets in the investment universe, risk parity portfolios allocate the total risk evenly across the assets.

**4.1.2.1. Constructing risk parity portfolios.** Risk parity portfolios have appealing properties that make them appropriate choices for investors looking for diversified portfolios. They are also easy to interpret and explain in intuitive terms. However, in comparison to equally-weighted or global minimum variance portfolios, they can be much harder to construct, especially in scenarios involving additional constraints on the final portfolio. One special case, namely the task of building a long-only risk-parity portfolio with no additional constraints is relatively straightforward as we explain below. However, it can be impossible to find a true risk-parity portfolio when the portfolio manager imposes discretionary constraints on the portfolio such as lower and upper bounds on individual component weights. Simple structural questions such as “is there always a risk parity portfolio associated with a given covariance matrix  $\Sigma$ ?”, or “when risk parity portfolios exist, are they unique?” may be hard to answer when coupled with constrained scenarios.

To pose the problem of finding a risk parity portfolio as an optimization problem, one first needs a quantification of the *deviation from risk parity*. Then, minimization of this quantity may produce a risk-parity portfolio. With this in mind, let us consider the following deviation measure proposed by Maillard, Roncalli, and Teiletche (2010)

$$\text{DRP}_1(\omega) = \sum_{i=1}^n \sum_{j=1}^n \omega_i (\Sigma \omega)_i - \omega_j (\Sigma \omega)_j^2$$

measuring all pair wise differences. Alternatively, we may choose to consider deviations from the average value in squared or absolute sense, i.e.

$$\text{DRP}_2(\omega) = \sum_{i=1}^n \left( \frac{\omega_i (\Sigma \omega)_i}{\omega^\top \Sigma \omega} - \frac{1}{n} \right)^2$$

$$\text{DRP}_3(\omega) = \sum_{i=1}^n \left| \frac{\omega_i (\Sigma \omega)_i}{\omega^\top \Sigma \omega} - \frac{1}{n} \right|$$

Now, the risk parity portfolio can be sought by minimizing any one of these deviation-from-parity functions. Unfortunately, all three measures we defined are non-convex functions of the portfolio weights. Therefore, the resulting minimization problem is likely to be difficult. Non-convexity of the deviation measure implies that the deviation measure may change unpredictably as one modifies the portfolio weights, and that the local minimizers of the deviation functions may not be global minimizers. Indeed, Maillard et al. (2010) report that the numerical optimization of function  $DRP_1(\omega)$  “is tricky” and they had to resort to heuristic approaches as outlined in their article.

As observed by Maillard et al. (2010) and Kaya and Lee (2012), an interesting and very useful characterization of risk-parity is possible in the long-only case. Let us consider the optimization problem

$$\min_{\omega > 0} f(\omega) = \omega^\top \Sigma \omega - \sum_{i=1}^n \ln \omega_i$$

The domain of the objective function  $f$  for this problem is the interior of the nonnegative orthant  $\mathbb{R}_+^n = \{x : x \geq 0\}$ . The optimization problem is unconstrained in its domain.

This optimization problem is sometimes referred to as the logarithmic penalty formulation since the term  $-\sum_{i=1}^n \ln x_i$  is the *logarithmic barrier* (or *logarithmic penalty*) function for the nonnegative orthant  $\mathbb{R}_+^n$ . The logarithmic barrier function is a well-known tool in optimization, especially in the context of interior-point methods used for solving constrained optimization problems (Boyd & Vandenberghe, 2004). The optimization problem we consider above with objective function  $f$  is related to but different from the formulations in Maillard et al. (2010) and Kaya and Lee (2012). Specifically, Maillard et al. (2010) minimize the risk term subject to a lower bound on the logarithmic barrier term, while Kaya and Lee (2012) maximize the logarithmic barrier term subject to a lower bound on the risk term. However, the three formulations are essentially equivalent and have very similar optimality conditions. The optimality conditions for the optimization problem above are given by

$$\nabla f(x) = 2\Sigma\omega - \omega^{-1} = 0$$

where  $\omega^{-1} = [\frac{1}{\omega_1}, \frac{1}{\omega_2}, \dots, \frac{1}{\omega_n}]^\top$  denotes the vector of reciprocals of the vector  $\omega$ . Note that we can write each element of the optimality conditions in the equivalent form

$$2(\Sigma\omega)_i - \frac{1}{\omega_i} = 0 \iff \omega_i \cdot (\Sigma\omega)_i = \frac{1}{2}, \quad i = 1, \dots, n$$

If these conditions hold, then we have

$$RRC_i(\omega) = \frac{\omega_i \cdot (\Sigma\omega)_i}{\omega^\top \Sigma \omega} = \frac{1/2}{n/2} = \frac{1}{n}$$

We have shown that the optimality conditions for the optimization problem are equivalent to the risk parity conditions given above. Therefore, optimal solutions of the problem we posed must be risk parity portfolios. There are other nice properties of this optimization problem in addition to the observation we just outlined. In particular, the objective function is strictly convex function of  $\omega$  which immediately implies that the solution is unique and is a global minimizer of the function  $f$ . We note that the weights  $\omega_i$  satisfying equation the optimality conditions may not add up to 1, so the solution is not necessarily a valid or desirable portfolio. However, it is easy to scale these weights so that the scaled weights satisfy a full investment constraint as well as the risk parity condition.

## 4.2. Mixing different sources of alpha

An important element of a quantitative portfolio selection process is the construction of a return model. The output from a return model is often expressed as a vector of “alpha’s”, i.e. a vector of forecasted excess returns for each one of the securities in the investment universe. Alphas represent the portfolio manager’s proprietary views on the investment choices and are the primary determinant of the portfolio’s performance. They are also the main quantities used to assess the manager’s skill.

While alpha is typically a single number representing the excess return forecast for the corresponding security it often combines a number of views and opinions. For example, multiple-factor return models are prominent examples where several different and potentially conflicting views are mixed together. Another example of the “view mixing” methodology is the approach of Black and Litterman (1992) discussed in Section 3.3.3.

While mixing views into a single alpha value is often practical, this approach is not always suitable or feasible. In such cases, it may be preferable to represent the different views as separate return models and address the view or model mixing problem at the optimization/portfolio construction stage. In the next subsections, we provide examples of portfolio construction scenarios where such formulations are used and how advanced optimization software and modeling can be used to solve these problems.

### 4.2.1. Combining strategic and tactical views

We consider a scenario with an investment universe of  $n$  securities  $S_1, S_2, \dots, S_n$  with uncertain future returns  $r_1, r_2, \dots, r_n$ . Let  $r = [r_1, \dots, r_n]^\top$  be the vector of these returns. As before, a portfolio is represented with the  $n$ -dimensional vector  $\omega = [\omega_1, \dots, \omega_n]^\top$  where  $\omega_i$  denotes the proportion of the total funds invested in security  $i$ . Let  $\Omega$ , a convex subset of  $\mathbb{R}^n$ , denote the set of permissible portfolios. The set  $\Omega$  is often defined through explicit constraints on the allocation of weights. We use this generic representation to simplify the exposition.

We assume that the portfolio is managed with respect to a benchmark portfolio  $\omega_B$ . The (uncertain) return of the portfolio,  $r_P$ , depends linearly on the weights

$$r_P(\omega) = \omega_1 r_1 + \dots + \omega_n r_n = \omega^\top r$$

As before,  $\Sigma$  represents the symmetric  $n \times n$  covariance matrix of the returns.

We consider two different return models for these securities. These models may differ on their return forecasts, for example, because of the different investment horizon assumptions they use. A typical example of this scenario is encountered when one tries to mix strategic (longer term) views with tactical (short-term) views. Let us represent the expected return coming from the strategic model with  $\mu^S$  and those coming from the tactical model with  $\mu^T$ ,

$$\mu^S = \begin{bmatrix} \mu_1^S \\ \vdots \\ \mu_n^S \end{bmatrix}, \quad \text{and} \quad \mu^T = \begin{bmatrix} \mu_1^T \\ \vdots \\ \mu_n^T \end{bmatrix}$$

Why would one not want to combine these two alphas vectors using some appropriate weighting parameters that address the difference between their horizon assumptions? There are several potential reasons:

- We may want to constrain risk contributions of each alpha source separately.
- We may want to separate out the active positions due to each alpha source for easier return and risk attribution.

- We may want to treat trades motivated by each alpha source differently as the expected holding period might be different.

To address these considerations, one needs to use two separate decision variables for each investment security in the mathematical formulation of the problem. Representing the active weights due to the strategic model with  $\omega_A^S$  and the active weights due to the tactical model with  $\omega_A^T$ , one may arrive at a formulation such as the following one

$$\begin{aligned} \max_{\omega \in \Omega} \quad & (\mu^S)^\top \omega_A^S + (\mu^T)^\top \omega_A^T - \lambda \cdot (\omega_A^S + \omega_A^T)^\top \Sigma (\omega_A^S + \omega_A^T) - \gamma \cdot TC(\omega) \\ \text{s.t.} \quad & \omega - \omega_A^S - \omega_A^T = \omega_B \\ & (\omega_A^S)^\top \Sigma \omega_A^S \leq u^S \\ & (\omega_A^T)^\top \Sigma \omega_A^T \leq u^T \end{aligned}$$

The separate risk constraints on the strategic and tactical active weights in the formulation above represent a *risk budgeting* approach where one makes a priori risk allocations to each view. The presence of such constraints prevents the portfolio risk to be dominated by one view or the other. These allocations may change over time depending on the performance of each view as well as the managers changing conviction level on the strategic and tactical views.

As we discussed in Section 3.1, the selection of the transaction cost aversion parameter  $\gamma$  is often linked to the average holding period of the investments in the portfolio as the one-time transaction cost of entering into or existing out of a position needs to be amortized over the life of holding the position. Representing the trading cost function  $TC(\omega)$  in terms of the aggregate weights rather than the individual weights due to strategic or tactical views obscures the potentially complicated treatment of the trading costs when one considers multiple views with multiple investment horizons.

On the one hand, the aggregation in the formulation above appears necessary. For example, if the strategic view is bullish on one security while the tactical view is bearish, optimal strategic and tactical weights may have opposite signs. This may require a buy trade for the strategic view and a sell trade for the tactical view. Of course, the portfolio manager would only execute the net trade rather than buying and selling the same security at the same time for the portfolio, justifying the aggregation as in the formulation above.

On the other hand, this may not be an adequate representation of the portfolio manager's problem if they expect the holding periods for strategic and tactical investments to be significantly different and instead wants to amortize the trading costs differently for trades motivated by each view by choosing different transaction cost aversion parameters for trades coming from each view. One could, therefore, replace the aggregate transaction cost term  $\gamma \cdot TC(\omega)$  with something like  $\gamma^S \cdot TC(\omega_A^S) + \gamma^T \cdot TC(\omega_A^T)$ . While this substitution would allow a differentiation of trading cost amortization, it is not clear how one would address net trades in this version of the formulation.

We think the proper treatment of different holding periods for different securities or views remains a relatively unexplored segment of the literature on quantitative portfolio construction and can benefit from further research. We discuss one possibility to address this issue in Section 4.3 by using multi-period optimization.

#### 4.2.2. Overlaying views on groups of securities

Consider the following scenario: A portfolio manager wants to build a global equity portfolio and has constructed a multi-factor model to forecast the future returns of the individual stocks in

the investment universe. In addition, the portfolio manager has built a macro-investing model that provides forecasts of returns for different countries represented in the equity investment universe. Here, the return for a country may be measured as the return to a broad index for that country's equity market. How can the portfolio manager combine country views with cross-sectional views on individual stocks? What should be done if, say, the portfolio manager is bullish on Toyota but bearish on Japan? Alternatively, consider the scenario where the portfolio manager wants to combine views on individual stocks with views on the industries or sectors those stocks belong to. If, say, the portfolio manager is bearish on technology stocks, but bullish on Apple, should this stock be overweighted or underweighted?

These are two very similar scenarios where two models are being mixed. The first model is on the individual elements in the investment universe consisting of securities  $S_1, S_2, \dots, S_n$ . The second model, however, is on groups of securities from this universe. In this section, we provide a possible mathematical formulation of this “overlay problem” and discuss methods for its solution.

A crucial element of these models is the mapping of securities to groups (countries, sectors, industries, etc.) as well as the associated hierarchy of decision variables. The security to group mapping can be described using the  $m \times n$  incidence matrix  $G$  where  $G_{gi} = 1$  if security  $i$  is in group  $g$ , and 0 otherwise. We also assume that there is a benchmark for the portfolio with given weights at the individual security level,  $\omega_B^I$ . In this paper, we use a superscript of  $I$  to describe decision variables or weight vectors expressed at the individual security level while the superscript of  $G$  is used with group level variables and quantities. We infer weights for a benchmark at the group level  $\omega_B^G$  using the equation  $\omega_B^G = G\omega_B^I$ .

There are several potential ways to model this particular model mixing problem. One of them is the following approach. We first consider the decisions made at the group level. Assuming that there are  $m$  groups representing the partition of the securities into countries, sectors, or industries, we denote the group weights with the vector  $\omega^G = [\omega_1^G, \dots, \omega_m^G]^\top$ .

The selections made at the group level can be used to generate a tilted/scaled benchmark at the security level constructed as follows. We preserve the relative weights of the securities in a group, but scale them up or down with respect to the original benchmark weights based on the ratio of the chosen group weight to the benchmark weight of the group. To be more specific, for each security in the benchmark, if it belongs to the group indexed by  $g$ , scale its weight by the ratio  $\omega_g^G / (\omega_B^G)_g$ .

This will produce a scaled benchmark where the weights add up to 1 as long as group weights add up to 1 and the group mapping represents a full partition of the investment universe. The scaled benchmark can be expressed mathematically by  $\hat{\omega}_B^I = \hat{G}\omega^G$  where  $\hat{G} = \text{diag}(\omega_B^I)G^\top \text{diag}(G\omega_B^I)^{-1}$ . Using this notation, we can express the active positions taken both at the group level as well as the individual security level

$$\omega_A^G = \omega^G - \omega_B^G$$

$$\omega_A^I = \omega^I - \hat{\omega}_B^I$$

Note that the active weights in the individual security level are defined with respect to the scaled benchmark rather than the actual benchmark. This allows the added value from the group level model and the security level model to be separated and attributed properly. Further, we do not require that  $G\omega^I = \omega^G$  which means that security level decisions do not need to conform to the group level decisions.

Using this notation, the full optimization problem takes the form



$$\begin{aligned}
& \max_{\omega^G \in \Omega^G, \omega^I \in \Omega^I} (\mu^G)^\top \omega_A^G + (\mu^I)^\top \omega_A^I - \lambda^G \cdot (\omega_A^G)^\top \Sigma^G \omega_A^G - \lambda^I \cdot (\omega_A^I)^\top \Sigma^I \omega_A^I - \gamma \cdot TC(\omega^G, \omega^I) \\
& \text{s.t. } \omega_A^G = \omega^G - \omega_B^G \\
& \quad \omega_A^I = \omega^I - \omega_B^I \\
& \quad (\omega_A^G)^\top \Sigma^G \omega_A^G \leq u^G \\
& \quad (\omega_A^I)^\top \Sigma^I \omega_A^I \leq u^I
\end{aligned}$$

#### 4.3. Multi-period optimization

The classical works of Merton (1969, 1990), and Campbell and Viceira (2002) illustrate that single-period portfolio choice policies are in general not optimal as they do not capture intertemporal effects and hedging demands. Return predictability and market impact naturally give rise to intertemporal hedging demands for securities, and investors need to look beyond just the next period when optimally allocating across securities. For instance, market impact costs from trades in the current period have an effect on prices in later periods.

Market impact is the effect that a portfolio manager has on prices from buying and selling assets. In general, prices move against the portfolio manager. Prices go up as a result of purchase; a sale results in a price decline. When he sells, prices go down. Market impact costs are costs associated with the immediacy of trading: In general, if the trading is done quickly (slowly) market impact costs are higher (lower). Portfolio managers frequently break up their orders (parent orders) into smaller pieces (child orders) and trade those over longer periods of time. However, if a portfolio manager is trading an order over a longer period of time, then due to price volatility there is a chance that asset prices adversely. In other words, there is a *trade-off* between (1) the cost of immediacy of trading (that comes with higher market impact costs), and (2) price volatility (that comes with lower market impact costs). So-called “optimal execution” software can model and optimize this trade-off associated with trading a given order (see, for example, Almgren and Chriss (2000)).

The objective of portfolio optimization is to find an optimal *trade-off between return and risk*. Traditionally, this was done *independently of the trade-off associated with trading* (described in the previous paragraph). However, this will often lead to target portfolio holdings that will incur large market impact costs. While not fully modeling market impact, a simple approach is to incorporate transaction costs into mean–variance optimization as we discussed in Section 3.1.

In this section we discuss a multi-period framework that allows us to jointly model risk, return predictability (alpha) and its decay, and impact costs. The models of Grinold (2006), Engle and Ferstenberg (2007), Garleanu and Pedersen (2013), and Kolm and Maclin (2012) are special cases of this model. A detailed description of this model, that also incorporates standard portfolio constraints, is presented in Kolm (2013).<sup>11</sup> Other recent work on multi-period optimization includes Boyd, Mueller, O'Donoghue and Wang (2013).

As before, we consider a discrete time economy with  $n$  risky securities. Suppose that price changes  $r_{t+1} \equiv p_{t+1} - p_t$  are given by

$$r_{t+1} = \mu_t + \alpha_t + \varepsilon_{t+1}^r$$

where  $\mu_t$  are the fair security returns (the return that correctly compensates for the risk of holding each security),  $\alpha_t$  are predictable excess returns, and  $\varepsilon_{t+1}^r$  is an unpredictable residual term with  $E_t(\varepsilon_{t+1}^r) = 0$  and  $\text{Var}_t(\varepsilon_{t+1}^r) = \Sigma$ . The investor forecasts “alphas” using a factor model with  $k$  mean reverting factors ( $k \ll n$ )

$$\alpha_t = Bf_t + \varepsilon_t^z$$

$$\Delta f_{t+1} = -Df_t + \varepsilon_{t+1}^f$$

In the first equation,  $f_t \in \mathbb{R}^k$  represents the factors,  $B \in \mathbb{R}^{n \times k}$  the factor loadings, and  $\varepsilon_t^f$  the idiosyncratic components. This specification generalizes a standard static factor model, making it dynamic and time-dependent. The second equation specifies the temporal behavior of the factors. Here  $D \in \mathbb{R}^{k \times k}$  is a positive definite matrix of mean-reversion coefficients. Intuitively, the greater the elements of this matrix the faster the factors mean revert to zero. Furthermore, we assume  $E_t(\varepsilon_{t+1}^z) = E_t(\varepsilon_{t+1}^f) = 0$ ,  $\text{Var}_t(\varepsilon_{t+1}^f) = \Sigma^f$ ,  $\text{Var}_t(\varepsilon_{t+1}^z) = \Sigma^z$ , and that the error terms  $\varepsilon_t^r, \varepsilon_t^z, \varepsilon_t^f$  are mutually independent. We will now incorporate temporary and permanent market impact costs into this model.

We denote the investor's holdings at time  $t$  by  $\omega_t$ . Any trading in our economy is subject to both temporary and permanent impacts. Following Almgren and Chriss (2000), we model the costs associated with trading an amount of  $\Delta\omega_t = \omega_t - \omega_{t-1}$  shares as a cost against the investor's alpha, that is

$$\begin{aligned}
\alpha_t &= Bf_t + \varepsilon_t^z + \underbrace{\Pi \Delta\omega_t}_{\text{permanent}} + \underbrace{H \Delta\omega_t}_{\text{temporary}} - H \Delta\omega_{t-1} \\
&= Bf_t + \varepsilon_t^z + \Pi \omega_t + H(\Delta\omega_t - \Delta\omega_{t-1})
\end{aligned}$$

where  $\Pi, H \in \mathbb{R}^{n \times n}$ . Note that the term  $H(\Delta\omega_t - \Delta\omega_{t-1})$  reverses the effect of a trade from one period to the next, making the impact of  $H \Delta\omega_t$  effective for one period only (single period impact). Therefore, we refer to  $H \Delta\omega_t$  as the temporary component and  $\Pi \Delta\omega_t$  as the permanent component.

In most practical applications temporary impacts do not decay instantaneously over a single period but rather last for several. We can achieve this by modeling them as

$$\alpha_t = Bf_t + \varepsilon_t^z + (\Pi + H)\Delta\omega_t - h_t$$

where the new state variable  $h_t$  (the remaining temporary impacts) is defined by

$$h_t = G h_{t-1} + (I - G)H \Delta\omega_{t-1} + \varepsilon_t^h$$

Here we assume  $E_t(\varepsilon_t^h) = 0$ , and  $\text{Var}_t(\varepsilon_t^h) = \Sigma^h$ . The matrix  $G \in \mathbb{R}^{n \times n}$  (with  $\|G\| < 1$ ) determines how fast the temporary impact decays.

The investor solves the following multi-period version of the mean–variance problem

$$\begin{aligned}
& \max_{\Delta\omega_1, \Delta\omega_2, \dots, \Delta\omega_{T-1}} E \left[ \left\{ \sum_{t=1}^{T-1} (1-\rho)^t \left( \omega_t^\top \alpha_t - \frac{\lambda}{2} \omega_t^\top \Sigma \omega_t - \frac{1}{2} \Delta\omega_t^\top A \Delta\omega_t \right) \right\} \right. \\
& \quad \left. + (1-\rho)^T \left( \omega_T^\top \alpha_T - \frac{\lambda}{2} \omega_T^\top \Sigma \omega_T \right) \right]
\end{aligned}$$

where  $\rho \in (0, 1)$  is a discount factor,  $\lambda$  is a risk aversion coefficient,  $\Sigma$  is the covariance matrix of returns,  $A$  is a diagonal matrix of positive entries representing a quadratic transaction cost, and  $\omega_0$  are the initial portfolio holdings. This objective function together with the dynamics turns out to be a stochastic linear-quadratic (LQ) regulator problem. In this model, the trade size  $\Delta\omega_t$  is the control variable.

We provide a sketch of the analytical details underlying this multi-period framework. Further details are presented in Kolm (2013). The problem can be formulated as a Linear–Quadratic–Gaussian

<sup>11</sup> See also Kolm and Ritter (2013) that introduce a novel approach for the study of general multi-period portfolio optimization problems with alphas, trading costs, and constraints. Their approach is inspired by the theory of Bayesian state estimation and is not limited to Linear–Quadratic–Gaussian models. They show that for a multi-period portfolio optimization problem, there is a “dual” Bayesian dynamic model where the most likely sequence of hidden states is the trading path which optimizes expected utility.

(LQG) control problem, where the state equations are linear and the cost function is quadratic. Its solution follows closely that of the standard theory (Åström, 1970).

First, let us rewrite the dynamics of the state variables ( $f_t$ ,  $w_t$ , and  $h_t$ ) in a linear state space representation. It is convenient to define the “augmented state” variable  $s_t = (f_t, w_t, h_t)'$ . Now, observe that its dynamics are given by

$$s_t = \begin{pmatrix} I-D & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & G \end{pmatrix} s_{t-1} + \begin{pmatrix} 0 \\ I \\ (I-G)H \end{pmatrix} \Delta\omega_t + \begin{pmatrix} e_t^f \\ 0 \\ e_t^h \end{pmatrix} = \hat{A}s_{t-1} + \hat{B}\Delta\omega_t + \varepsilon_t$$

where

$$\text{Var}_{t-1}(\varepsilon_t) = \begin{pmatrix} \Sigma_f & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Sigma_h \end{pmatrix}$$

Now, we observe that each summand of the multi-period objective function is quadratic and can be rewritten as follows:

$$\begin{aligned} \omega_t^\top \alpha_t - \frac{\lambda}{2} \omega_t^\top \Sigma \omega_t - \frac{1}{2} \Delta\omega_t^\top A \Delta\omega_t &= \omega_t^\top (Bf_t + e_t^z + (\Pi + H)\Delta\omega_t - h_t) \\ &- \frac{\lambda}{2} \omega_t^\top \Sigma \omega_t - \frac{1}{2} \Delta\omega_t^\top A \Delta\omega_t = \begin{pmatrix} s_t \\ \Delta\omega_t \end{pmatrix}^\top \begin{pmatrix} R & S \\ S' & Q \end{pmatrix} \begin{pmatrix} s_t \\ \Delta\omega_t \end{pmatrix} \\ &+ O(e_t^z) \equiv c(s_t, \Delta\omega_t) + O(e_t^z) \end{aligned}$$

where

$$R = \begin{pmatrix} 0 & \frac{1}{2}B^\top & 0 \\ \frac{1}{2}B & -\frac{\lambda}{2}\Sigma & -\frac{1}{2}I \\ 0 & -\frac{1}{2}I & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 0 \\ \frac{1}{2}(\Pi + H) \\ 0 \end{pmatrix}, \quad \text{and} \quad Q = -\frac{1}{2}A$$

Together, the objective function therefore takes the form

$$\min_{\Delta\omega_1, \Delta\omega_2, \dots, \Delta\omega_{T-1}} E \left[ \sum_{t=1}^{T-1} (1-\rho)^t c(s_t, \Delta\omega_t) + (1-\rho)^T C(s_T) \right]$$

where  $c(s, \Delta\omega)$  is defined above,  $C(s) = s^\top R s$ , and  $R = R^\top \geq 0$ ,  $S \geq 0$ ,  $Q = Q^\top > 0$ .

The linear dynamics and the quadratic objective function define an LQG problem. Applying standard results from the theory of LQG problems the optimal control is the linear control given by  $\Delta x_t = L_t s_t$ . Here the matrix  $L_t$  is given by

$$L_t = -(Q + (1-\rho)\hat{B}^\top K_{t+1} \hat{B})^{-1} (S + (1-\rho)\hat{B}^\top K_{t+1} \hat{A}), \quad t < T$$

where  $K_t$  satisfies the Riccati equation

$$K_t = R + (1-\rho)\hat{A}^\top K_{t+1} \hat{A} - (S^\top + (1-\rho)\hat{A}^\top K_{t+1} \hat{B})(Q + (1-\rho)\hat{B}^\top K_{t+1} \hat{B})^{-1} S + (1-\rho)\hat{B}^\top K_{t+1} \hat{A}$$

The calculations of these control matrices are straightforward and only involve matrix operations. A full analysis of this model is beyond the scope of this paper, but we discuss the intuition behind the main results here.

As one would expect, the model prefers securities with high alphas, and securities that decay slowly relative to other securities. Of course, these are the assets with higher loadings to the more persistent factors. If either alpha decreases or its decay increases, the resulting holdings decrease, and vice versa. We refer to the optimal portfolios determined without trading costs (for each  $t$ ) as the *target portfolios*.

In the sole presence of one-period temporary market impact cost, the optimal trade is proportional to the difference between the current and target portfolio holdings. When temporary market impact costs last for more than one period – and/or in the presence of permanent market impact costs – the optimal trade is

proportional to the difference between current portfolio holdings and a dynamically modified optimal target portfolio (from here on referred to as the *dynamic target portfolio*). This dynamic target portfolio is different from the optimal target portfolio above, as it incorporates the transient and persistent frictional effects of transaction costs from trading the securities.

We point out that dynamic portfolio analysis changes the perception of traditional portfolio management – from a static to a dynamic view. Portfolios are moving targets that we are attempting to optimally track in order to balance the trade-off between risk-adjusted returns and trading costs while accounting for the persistent effects of both.

In practice, multi-period models are seldom used. There are several practical reasons for that. First, it is often very difficult to accurately estimate return/risk for multiple periods, let alone for a single period. Second, multi-period models are in general computationally intensive, especially if the universe of assets considered is large. Third, the most common existing multi-period models do not handle real-world constraints. As discussed in Section 3.2, in practice the majority of managed portfolios are subject to all kinds of constraints. For these reasons, practitioners typically use single-period models to rebalance the portfolio from one period to another.

The concepts of portfolio optimization and diversification have been instrumental in the development and understanding of financial markets and financial decision making. In light of the 60 year anniversary of Harry Markowitz's paper “Portfolio Selection,” in the previous sections of this paper we reviewed some of the approaches developed to address the challenges encountered when using portfolio optimization in practice, including the inclusion of transaction costs, portfolio management constraints, and the sensitivity to the estimates of expected returns and covariances. In addition, we selectively highlighted some of the new trends and developments in the area such as diversification methods, risk-parity portfolios, the mixing of different sources of alpha, and practical multi-period portfolio optimization. This will without doubt continue to be an important area both in academic theory as well as practical settings.

## Acknowledgements

We are grateful for helpful comments from Joseph Cerniglia as well as the participants of the conference “Celebrating the 60th Anniversary of Modern Portfolio Theory” honoring Nobel Laureate Harry Markowitz held on November 27, 2012, at Fordham University in New York.

## References

- Almgren, R., & Chriss, N. (2000). Optimal execution of portfolio transactions. *Journal of Risk*, 3, 5–39.
- Almgren, R., Thum, C., Hauptmann, E., & Li, H. (2005). Equity market impact. *Risk*, 18, 57–62.
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 3, 203–228.
- Asness, C., Frazzini, A., & Pedersen, L. H. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, 68, 47–59.
- Åström, K. J. (1970). *Introduction to stochastic control theory*. Academic Pr.
- Athayde, G. M., & Flöres, R. G. Jr., (2002). The portfolio frontier with higher moments: The undiscovered country. *Computing in Economics and Finance, Society for Computational Economics*.
- Athayde, G. M., & Flöres, R. G. Jr., (2004). Finding a maximum skewness portfolio – A general solution to three-moments portfolio choice. *Journal of Economic Dynamics and Control*, 28, 1335–1352.
- Bender, J., Lee, J. -H., & Stefek, D. (2009). Decomposing the impact of portfolio constraints. In *MSCI barra research insights: MSCI Barra*.
- Ben-Tal, A., & Nemirovski, A. S. (1998). Robust convex optimization. *Mathematics of Operations Research*, 23, 769–805.
- Ben-Tal, A., & Nemirovski, A. S. (1999). Robust solutions to uncertain linear programs. *Operations Research Letters*, 25, 1–13.

- Best, M., & Grauer, R. (1991a). The analytics of sensitivity analysis for mean variance portfolio problems. *International Review of Financial Studies*, 4, 315–342.
- Best, M., & Grauer, R. (1991b). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *Review of Financial Studies*, 4, 315–342.
- Black, F., & Litterman, R. (1991). Asset equilibrium: Combining investor views with market equilibrium. *Journal of Fixed Income*, 1, 7–18.
- Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48, 28–43.
- Bonnans, J. F., & Shapiro, A. (2000). *Perturbation analysis of optimization problems*. New York: Springer-Verlag.
- Borkovec, M., Domowitz, I., Kiernan, B., & Serbin, V. (2010). Portfolio optimization and the cost of trading. *Journal of Investing*, 19, 63–76.
- Bouchard, J. -P., Potters, M., & Aguilar, J. -P. (1997). Missing information and asset allocation. In *Science & finance: Capital fund management*.
- Boyd, S., Mueller, M., O'Donoghue, B., & Wang, Y. (2013). Performance bounds and suboptimal policies for multi-period investment. *Foundations and Trends in Optimization*, 1, 1–69.
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge: Cambridge University Press.
- Brandes, Y., Domowitz, I., & Serbin, V. (2012). Transaction costs and equity portfolio capacity analysis. In B. Scherer, & K. Winston (Eds.), *The oxford handbook of quantitative asset management*. OUP Oxford.
- Brown, D. B., & Smith, J. E. (2011). Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. *Management Science*, 57, 1752–1770.
- Campbell, J. Y., & Viceira, L. M. (2002). *Strategic asset allocation: Portfolio choice for long-term investors*. USA: Oxford University Press.
- Ceria, Takriti, Tierens, & Sofianos (2008). Incorporating the Goldman Sachs shortfall model into portfolio optimization. In *Axioma's breakfast research seminar series*. New York.
- Ceria, S., Saxena, A., & Stubbs, R. A. (2012). Factor alignment problems and quantitative portfolio management. *Journal of Portfolio Management*, 38, 29–43.
- Ceria, S. (2012). To optimize or not to optimize: Is that the question? In B. Scherer & K. Winston (Eds.), *The oxford handbook of quantitative asset management* (pp. 50–64). New York: Oxford University Press.
- Chen, A., Fabozzi, F. J., & Huang, D. (2010). Models for portfolio revision with transaction costs in the mean-variance framework. In J. B. Guerard, Jr. (Ed.), *The handbook of portfolio construction: Contemporary applications of markowitz techniques*. New York: Springer.
- Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19, 6–11.
- Clarke, R., De Silva, H., & Thorley, S. (2002). Portfolio constraints and the fundamental law of active management. *Financial Analysts Journal*, 58, 48–66.
- Cornuejols, G., & Tütüncü, R. (2007). *Optimization methods in finance*. Cambridge, UK: Cambridge University Press.
- Cremers, J. -H., Kritzman, M., & Page, S. (2003). *Portfolio formation with higher moments and plausible utility*. 272-12 Revere Street Working Papers.
- Cremers, J. -H., Kritzman, M., & Page, S. (2005). Optimal hedge fund allocations – Do higher moments matter? *Journal of Portfolio Management*, 3, 70–81.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22, 1915–1953.
- DeMiguel, V., & Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research*, 57, 560–577.
- El Ghaoui, L., & Lebret, H. (1997). Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18, 1035–1064.
- El Ghaoui, L., Oks, M., & Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51, 543–556.
- El Ghaoui, L., Oustry, F., & Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9, 33–52.
- Ellis, K., Michaely, R., & O'Hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis*, 35, 529–551.
- Engle, R. F., & Ferstenberg, R. (2007). *Execution Risk Journal of Portfolio Management*, 33, 34–44.
- Fabozzi, F. J., Focardi, S. M., & Kolm, P. N. (2010). *Quantitative equity investing: Techniques and strategies*. Wiley.
- Fabozzi, F. J., Huang, D., & Zhou, G. (2010). Robust portfolios: Some recent contributions from operations research and finance. *Annals of Operations Research*, 176, 191–220.
- Fabozzi, F. J., Kolm, P. N., Pachamanova, D., & Focardi, S. M. (2007). *Robust portfolio optimization and management*. New York: Wiley.
- Frost, P. A., & Savarino, J. E. (1986). An empirical Bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis*, 21, 293–305.
- Garlappi, L., Uppal, R., & Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20, 41–81.
- Garleanu, N. B., & Pedersen, L. H. (2013). Dynamic trading with predictable returns and transaction costs. *Journal of Finance*, 68, 2309–2340.
- Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28, 1–38.
- Green, R. C., & Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified? *Journal of Finance*, 47, 1785–1809.
- Gregory, C., Darby-Dowman, K., & Mitra, G. (2011). Robust optimization and portfolio selection: The cost of robustness. *European Journal of Operational Research*, 212, 417–428.
- Grinold, R. (2005). Implementation efficiency. *Financial Analysts Journal*, 61, 52–64.
- Grinold, R. (2006). A dynamic model of portfolio management. *Journal of Investment Management*, 4, 5–22.
- Gupta, F., & Eichhorn, D. (1998). Mean-variance optimization for practitioners of asset allocation. In *Handbook of portfolio management*. New Hope, Pennsylvania: Frank J. Fabozzi Associates.
- Harvey, C. R., Liechty, J. C., Liechty, M. W., & Mueller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10, 469–485.
- Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, 55, 1263–1295.
- Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance*, 46, 179–207.
- J. P. Morgan & Co. (1997). *Creditmetrics – Technical document*.
- Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58, 1651–1684.
- Jean, W. H. (1971). The extension of portfolio analysis to three or more parameters. *Journal of Financial and Quantitative Analysis*, 6, 505–515.
- Jobson, J., & Korkie, B. (1981). Putting markowitz theory to work. *Journal of Portfolio Management*, 7, 70–74.
- Jobst, N. J., & Zenios, S. A. (2001). The tail that wags the dog: Integrating credit risk in asset portfolios. *Journal of Risk Finance*, 3, 31–44.
- Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business*, 58, 259–278.
- Jorion, P. (1986). Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21, 279–292.
- Jorion, P. (1991). Bayesian and CAPM estimators of the means: Implications for portfolio selection. *Journal of Banking and Finance*, 15, 717–727.
- Kallberg, J. G., & Ziemba, W. T. (1984). Misspecification in portfolio selection problems. *Risk and Capital*, 227, 74–87.
- Kaya, H., & Lee, W. (2012). *Demystifying risk parity*. Working paper.
- Kim, W. C., Kim, J. H., & Fabozzi, F. J. (2012). Deciphering robust portfolios. In *Working paper: Working paper*.
- Kim, J. H., Kim, W. C., & Fabozzi, F. J. (in press). Recent developments in robust portfolios with a worst-case approach. *Journal of Optimization Theory and Application*.
- Kim, W. C., Kim, M. J., Kim, J. H., & Fabozzi, F. J. (2014). Robust portfolios that do not tilt factor exposure. *European Journal of Operational Research*, 234, 411–421.
- Kim, W. C., Kim, J. H., Ahn, S. H., & Fabozzi, F. J. (2012). What do robust equity portfolio models really do? *Annals of Operations Research*, 205, 141–168.
- Kim, J. H., Kim, W. C., & Fabozzi, F. J. (2013). Composition of robust equity portfolios. *Finance Research Letters*.
- Kolm, P. N. (2013). *Multi-period portfolio optimization with market impact, alpha decay, and constraints*. Working paper, Courant Institute of Mathematical Sciences, New York University.
- Kolm, P. N., & Maclin, L. (2012). Algorithmic trading, optimal execution, and dynamic portfolios. In B. Scherer, & K. Winston (Eds.), *The oxford handbook of quantitative asset management*. OUP Oxford.
- Kolm, P. N., & Ritter, G. (2013). *Multiperiod portfolio choice and bayesian dynamic models*. Working paper, Courant Institute of Mathematical Sciences, New York University.
- Kritzman, M. (2006). Are optimizers error maximizers? *Journal of Portfolio Management*, 32, 66–69.
- Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10, 603–621.
- Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30, 110–119.
- Lee, C. M. C., & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance*, 46, 733–746.
- Lee, J.-H., & Stefek, D. (2008). Do risk factors eat alphas? *Journal of Portfolio Management*, 34, 12–25.
- Levy, H., & Markowitz, H. M. (1979). Approximating expected utility by a function of mean and variance. *American Economic Review*, 69, 308–317.
- Lillo, F., Farmer, J. D., & Mantegna, R. N. (2003). Master curve for price-impact function. *Nature*, 421, 129–130.
- Litterman, R. (2003). *Modern investment management*. Hoboken, New Jersey: John Wiley & Sons Inc..
- Lobo, M. S., & Boyd, S. (2000). *The worst-case risk of a portfolio*. Technical report.
- Maillard, S., Roncalli, T., & Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *Journal of Portfolio Management*, 36, 60–70.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7, 77–91.
- Markowitz, H. M. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Cambridge, MA: Basil Blackwell.
- Markowitz, H. (2014). Mean-variance approximations to expected utility. *European Journal of Operational Research*, 234, 346–355.
- Merton, R. C. (1969). lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51, 247–257.
- Merton, R. C. (1990). *Continuous-time finance*. Blackwell Publishers.
- Meucci, A. (2008). Fully flexible views: Theory and practice. *Fully Flexible Views: Theory and Practice*, Risk, 21, 97–102.
- Meucci, A. (2010). In Cont, R., & Tankov, P. (Eds.), *The Black-Litterman approach: Original model and extensions in the encyclopedia of quantitative finance*. Hoboken, NJ: John Wiley & Sons. Inc.

- Michaud, R. (1998). *Efficient asset allocation: A practical guide to stock portfolio optimization and asset allocation*. Boston: Harvard Business School Press.
- Morgan, J. P. (1996). *RiskMetrics™ – Technical document* (4th ed.). New York: Morgan Guaranty Trust Company of New York.
- Qian, E., & Gorman, S. (2001). Conditional distribution in portfolio theory. *Financial Analysts Journal*, 57, 44–51.
- Rockafeller, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21–41.
- Rubinstein, M. (2002). Markowitz's "portfolio selection": A fifty-year retrospective. *Journal of Finance*, 57, 1041–1045.
- Satchel, S., & Scowcroft, A. (2000). A demystification of the Black–Litterman model: Managing quantitative and traditional portfolio construction. *Journal of Asset Management*, 1, 138–150.
- Saxena, A., Martin, C., & Stubbs, R. A. (2012). *Aligning alpha and risk factors, a panacea to factor alignment problems?* Axioma research paper.
- Scherer, B. (2007). Can robust portfolio optimisation help to build better portfolios? *Journal of Asset Management*, 7, 374–387.
- Scherer, B., & Xu, X. (2007). The impact of constraints on value-added. *Journal of Portfolio Management*, 33, 45–54.
- Stein, D. M. (2001). Equity portfolio structure and design in the presence of taxes. *Journal of Wealth Management*, 4, 37–42.
- Stein, D. M., & Garland, J. P. (2008). Investment management for taxable investors. In F. J. Fabozzi (Ed.), *Handbook of finance* (Vol. II). Wiley.
- Steinbach, M. C. (2001). Markowitz revisited: Mean–variance models in financial portfolio analysis. *SIAM Review*, 43, 31–85.
- Stubbs, R. A., & Vandenbussche, D. (2010). Constraint attribution. *Journal of Portfolio Management*, 36, 48–59.
- Trojani, F., & Vanini, P. (2002). A note on robustness in Merton's model of intertemporal consumption and portfolio choice. *Journal of Economic Dynamics and Control*, 26, 423–435.
- Tütüncü, R. H., & Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132, 157–187.
- Tütüncü, R. (2012). Recent advances in portfolio optimization. In B. Scherer & K. Winston (Eds.), *The oxford handbook of quantitative asset management* (pp. 7–31). Oxford University Press.
- Zhu, S., & Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57, 1155–1168.