PRACTICAL PORTFOLIO OPTIMIZATION

Isaac Siwale

Abstract

This paper is on the portfolio optimization problem for which two generic models are presented in the context of a proprietary solver called GENO:[1] the first is a pseudo-dynamic model with a single state variable that is meant for the single holding-period case; the second is a truly dynamic model with several state variables that applies to both single and multi-period scenarios. Both models can handle practical restrictions such as exposure, cardinality and round-lot constraints; both can accommodate non-traditional risk measures; and the usual two-element criterion set comprising 'portfolio reward' and 'portfolio risk' may be augmented by any number of sub-objectives deemed necessary. The paper also presents a 'robust optimization' model of the portfolio problem that explicitly accounts for data uncertainty; the robust model can also accommodate sophisticated risk measures, and one such function is presented in **Appendix B**. The compromise solution concept is used to compute portfolios that are Pareto-efficient; several numerical examples show that the portfolios thus found are not only optimal in the compromise sense, but they also have a competitive (and often the highest) reward-risk ratio—a proxy for the Sharpe ratio. A limited empirical analysis shows that the robust model is superior to its non-robust counterpart in terms of both nominal performance, and the potential for avoiding "opportunity costs" due data uncertainty.

Key Words: Portfolio Selection; Robust Optimization; Multi-objective Programming; Evolutionary Algorithms; Compromise Solution; Coherent Risk Measures; Interval Programming.

1 Introduction

The portfolio problem first formulated by Harry M. Markowitz [42] is concerned with how investors ought to deploy their wealth in financial assets, and his recommendation is a strategy he called the 'E-V Rule':

> "The investor should select one of those portfolios with minimum *variance* [V] for a given *expected return* [E] or more, and maximum expected return for given variance or less" [Paraphrased from 42, p.82; emphasis added]

One may formally state Markowitz' normative decision model in terms of utility theory as follows:[2] assume an investor has a (subjective) utility function U, and there are n securities with random single-period returns r_i and covariance matrix is $[c_{ij}]$ in which he wishes to invest; then the best wealth deployment strategy may be found by solving the following mathematical program:[3]

MP1:
$$\underset{u}{\text{Max}}\ \mathbb{E}\left\{ U\!\left(w_0, r_i, u_i\right) \,\middle|\, \sum_i u_i \le w_0;\ u_i \ge 0 \right\}$$

Where: $\mathbb{E}$ — is the mathematical expectation operator;

               w_0 — is the investor's initial wealth;

               u_i — is the proportion of wealth allocated to the *i*-th asset;

               r_i — is the single-period rate of return of the *i*-th asset.

---

[1] This is an acronym for: *G*eneral *E*volutionary *N*umerical *O*ptimizer; the GENO solver is described in detail in [61].

[2] The seminal text for utility theory is [72], but a very readable account may also be found in [39].

[3] The abbreviation 'MP' or the contraction 'M-program' shall often be used in place of 'mathematical program' throughout this paper.

MP1 is a stochastic optimization problem, and one could use a stochastic algorithm to solve it. But, apart from the perils associated with calibrating utility functions (see Remarks in **Appendix A**), stochastic algorithms *per se* are computationally demanding and rather cumbersome. And in addition, their solution can be misleading when there is ambiguity in the choice of a distribution for the random parameters. And so, although modern integrated stochastic programming environments such as 'SPInE' [71] do alleviate some of the difficulties, the fact still remains that such a direct approach would be inefficient. And it was his recognition of such computational difficulties at that time, in conjunction with his quest to reconcile the E-V Rule with utility theory that prompted Markowitz to suggest a different approach; later, he explained his strategy as follows:

> "If an investor with a particular single period utility function acted only on the basis of expected return and variance, could the investor achieve almost maximum expected utility? Alternatively, if one knew the expected value and variance of a probability distribution of return on a portfolio can one guess fairly closely its expected utility?" [Paraphrased from 41, p.471]

Markowitz took a pragmatic approach to this question and derived a deterministic function that one could use in place of the utility function. As shown in **Appendix A**, his mean-variance (MV) criterion is, in essence, a Taylor series approximation of a general utility function. Empirical evidence shows that the MV criterion is quite representative of a variety of standard utility functions [1, 23, 35, 38, 53]; furthermore, although the MV criterion is only a second-order approximation, it suffices in most cases, i.e. no advantage is gained by retaining higher order terms in the Taylor series [29]. For most practical purposes therefore, the stochastic objective function in MP1 is normally replaced by the MV criterion and the resulting decision model is a two-term, uni-objective optimization problem that includes a measure of the model-user's tolerance for risk, viz.:

MP2: $$\underset{\mathbf{u}}{\text{Max}}\left\{ \sum_{i=1}^{n}\left(u_i r_i - \varphi^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} u_i u_j c_{ij} \right) \,\middle|\, \sum_i u_i = 1;\, u_i \geq 0 \right\}$$

Where: r_i — is the expected return for the *i*-th asset;

c_{ij} — is the covariance of returns between the *i*-th and *j*-th assets;

u_i — is the proportion of wealth invested in the *i*-th asset;

φ — is the investor's risk tolerance.

And in accordance with the EV rule, the MV model is usually stated and studied in one of two ways: (i) as MP3a wherein the portfolio return is required to equal or surpass an arbitrarily fixed target return, $\bar{r}$; (ii) as MP3b wherein the portfolio variance is required to be within an arbitrarily fixed target variance, γ. Research is focused more on MP3a (sometimes without the 'minimum-return' constraint) because solutions computed by this 'minimum variance' model are more stable with respect to perturbations in the input data [12].

MP3a: $$\underset{\mathbf{u}}{\text{Min}}\left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} u_i u_j c_{ij} \,\middle|\, \sum_i u_i r_i \geq \bar{r};\, \sum_i u_i = 1;\, u_i \geq 0 \right\}$$

MP3b: $$\underset{\mathbf{u}}{\text{Max}}\left\{ \sum_{i=1}^{n} u_i r_i \,\middle|\, \sum_{i=1}^{n} \sum_{j=1}^{n} u_i u_j c_{ij} \leq \gamma;\, \sum_i u_i = 1;\, u_i \geq 0 \right\}$$

This paper contributes a solution method whose main attributes are simplicity, practicality, and the fact that the multi-objective essence of the portfolio optimization is not suppressed. The paper is organized as follows: §**2** discusses practical issues in portfolio selection; §**3** presents two basic decision models for the problem; §**4** discusses special constraints; §**5** presents a method for handling uncertainty; the solution to the models is explained in §**6**; numerical examples are presented in §**7** and discussed in §**8**; §**9** summarises and concludes the presentation; last but not least, the legal framework governing this publication is set forth in §**10**.

2 Practical Issues

Preamble. The MV model is currently the standard normative theory for investor behaviour; its diversification principles are the foundation of 'Modern Portfolio Theory', and it is the basis for many other important advances in financial economics, including the celebrated Capital Asset Pricing Model [59]. However, Richard Michaud notes that:

> "There is some evidence that a number of experienced investment professionals have experimented with MV optimizers only to abandon the effort when they found their MV portfolios to be unintuitive and without obvious investment value. Given the success of the MV model as a conceptual framework, and the availability for nearly 30 years of [the Critical Line algorithm[4]] for computing efficient frontiers, it remains one of the outstanding puzzles of modern finance that MV optimization has yet to meet with widespread acceptance by the investment community [. . .] Does this 'Markowitz optimization enigma' reflect a considered judgement by the investment community that such methods are not worthwhile?" [Paraphrased from 46, p.31]

The evidence mentioned by Michaud may be anecdotal but it is nonetheless real and the 'enigma' still persists more than 20 years on; Kolm [34, p.7] quotes a modern senior portfolio manager at a large hedge fund thus:

> "We don't use MVO, Black-Litterman or any other methodology. We always combine signals using home grown techniques that we develop to suit our needs. The same holds for portfolio construction: There are a lot of competing objectives in play that make it hard to cast as a classical optimization problem. That's why we ended up with a lot of techniques that look ad-hoc, but are robust and work well for us. They are probably not optimal in a mathematical sense, but personally I consider optimality to be ill-defined in a noisy system like the stock market. We never spend a lot of time trying to make things optimal. Robustness is far more important to us."

Michaud's Diagnosis. The reluctance on the part of some practitioners to use the MV model is intriguing. Michaud [46] has examined this 'Markowitz optimization enigma' and his diagnosis is multi-faceted:

- Error Maximization. The unintuitive character of many "optimized" MV-portfolios can be traced to the fact that MV-optimizers are fundamentally "estimation-error maximizers". The data required for MV optimization are necessarily estimates that are subject to estimation errors and the optimization process significantly over-weights (under-weights) those securities that have large (small) estimated returns, negative (positive) correlations and small (large) variances. These securities are the ones most likely to have large estimation errors [9, 46].

- Instability of Solutions. MV optimizations can be very unstable whereby small changes in input data assumptions leads to large changes in the solution. The main reason for this is lies in the drastic loss of statistical efficiency (see footnote **8**) that some estimators suffer when the 'Gaussian error' assumption is violated [12, 46, 70].

- Non-Investability of Allocations. Standard MV Optimizers often produce portfolios that are not "investable". Because most Optimizers assume continuous variables, they can sometimes produce portfolio allocations that are not in convenient units, or are not in significant enough quantities, or both. And yet it is known that institutional assets managers typically make investments in large dollar increments, or in round lots of stocks [46].

- Special Constraints. Standard MV Optimizers ignore some important constraints that apply in practice. For example, under German Investment Law, there exists law—called the '5-10-40 rule'— that constrains the composition of mutual fund portfolios. This rule states: "up to 5% of the value of the mutual fund may be invested in securities and money market instruments of the same issuer. This limit may be extended to 10% if the contractual terms and conditions of the investment fund so provide and, at the same time, the total investment in the assets of such issuers does not exceed 40% of the net asset value of the fund" [22, p.11].

The MV model depends on forecasts of future events and as such it inherently entails "opportunity costs"; these are the potential costs of holding a supposedly optimal portfolio based as it is on a particular forecast of asset returns and their correlation when the said forecasts do not actually pan out as expected. Since this is usually the case in practice—because we can never know the future with absolute certainty—this paper proposes that a portfolio's "opportunity cost" should be part of its evaluation. And so, whilst acknowledging Michaud's 'Instability' diagnosis, 'portfolio robustness' shall be treated in this sense.

---

[4] For a recent review and typical application of the Critical Line algorithm, see [51].

Roy's Critique. There is also the question of whether the MV model realistically encapsulates investor behaviour. The model is normative, and as such, one may expect that clients would only have confidence in its prescriptions if these reflect common sense. Roy [57] questioned the fundamental premise of the MV model, namely that investors should maximize their expected utility: he did not believe that an accurate utility function could be derived for any investor, and therefore investors would find it practically impossible to "maximise their expected utility". Instead, Roy posited that investors normally decide using a strategy he calls 'Safety of Principal First'—they set some minimum acceptable return (dabbed the 'disaster level'), and their preferred portfolio is that whose return is least likely to fall below the disaster level, i.e. that which exhibits the least *downside* risk. Nawrocki observes that Markowitz himself was an early proponent of a similar idea:

> "[Markowitz] realized that (a) only downside risk is relevant to investors, and (b) security distributions may not be Gaussian distributed, in which case the use of variance is incorrect. In 1959, he suggested two downside risk measures—a semi-variance computed from the mean return, and a semi-variance computed from a target return. But Markowitz stayed with variance because it was simpler to compute" [Paraphrased from 50, p.3]

Portfolio Risk. Roy's positive theory has since spawned a spate of research into downside risk and its various the measures: a brief outline of the major strands of this research is as follows:

- Lower Partial Moment — The Lower Partial Moment of the random variable Z is defined as follows:

$$\mathbf{LPM}(\tau) = \mathrm{E}\left\{ \left(\mathrm{Max}\{\tau - Z; 0\} \right)^{\gamma} \right\}$$

  LPM subsumes "a significant number of the known Von Neumann-Morganstern utility functions, as well as the whole gamut of human behaviour from risk-seeking to risk-neutral to risk-averse" [50]. The range parameter τ is set equal to the threshold from which the deviations are calculated. The (subjective) parameter γ represents the investor's risk attitude—risk-averse investors are expected to choose a γ coefficient greater than 1 (giving more weight to large deviations); risk-neutral investors are expected to choose a γ equal to 1; and the risk-seekers would choose a γ that is less than 1; for more details, see e.g. [33].

- Value-at-Risk — In 1995 the Investment Bank, JP Morgan, released data on variances and covariances across various security and asset classes that it had used internally to manage risk; it called the service 'RiskMetrics' and used the term 'Value at Risk' to describe the risk measure that emerged from these data. This term (usually abbreviated as 'VaR') refers to a risk notion that addresses the question: 'What is the most I could lose on this investment?' VaR is formally defined as follows: suppose the random variable Z has an elliptical distribution (see footnote **5**) denoted by F_Z; then the Value-at-Risk of Z at probability level α is given by:

$$\mathrm{VaR}_{\alpha}(Z) = F_Z^{-1}(\alpha) = \inf\{Z \in \mathbb{R} : F_Z(Z) \geq \alpha\}$$

  In other words, VaR is simply an alternative notation for the quantile function of the distribution, F_Z, evaluated at probability level α. And up until the development of the theory of coherent risk measures (see below) where VaR was shown not to conform to this important idea in general, VaR was the de facto standard measure of risk in financial services firms and was beginning to find acceptance in the non-financial sector as well. For a historical review of VaR, including the associated technical details, see [47].

- Coherency — This term applies to risk measures that satisfy a set of axioms first articulated by Artzner, Delbaen, Eber & Heath [3, 4]; a very readable account of 'coherency' that also extends their formulation may be found in [56]. The Theory of Coherent Risk Measures—developed in consultation with risk-management practitioners—is an attempt to provide objective criteria for evaluating different quantitative measures of risk. A coherent risk measure is a real-valued function ρ on the space of real-valued random variables that exhibits:

  A1. *Monotonicity*: For any two random variables, if $Z_1 \leq Z_2$, then $\rho(Z_2) \leq \rho(Z_1)$

  A2. *Positive Homogeneity*: For $\lambda \geq 0$ we have that $\rho(\lambda Z) = \lambda \rho(Z)$

  A3. *Translation Invariance*: For any $\alpha \in \mathbb{R}$ we have that $\rho(Z + \alpha) = \rho(Z) - \alpha$

  A4. *Sub-additivity*: For any two random variables Z_1 and Z_2 we have $\rho(Z_1 + Z_2) \leq \rho(Z_1) + \rho(Z_2)$

These properties are consistent with common sense: A1 implies that if portfolio Z_2 is preferred to portfolio Z_1, then Z_1 is more risky than Z_2; A2 reflects the fact that risk "grows" in direct proportion to the size of the portfolio; A3 means that if a component with deterministic outcome is added to the current holding, the risk of the new portfolio is reduced by the same amount; and A4 essentially states that the risk of a portfolio cannot exceed the sum of risks of its individual components, in other words, "a merger does not create extra risk" [4]. In the portfolio selection context, A4 is the most important property—it is what makes diversification beneficial.

The case for coherent downside risk measures in portfolio optimization is incontestable. In an ideal world where portfolio returns are elliptically distributed, this requirement would not be so crucial because, in that scenario, most risk measures would satisfy what Embrechts, McNeil & Straumann [16, 17] call the "Fundamental Theorem of Integrated Risk Management", which is essentially a statement on the suitability of risk measures under various probabilistic assumptions; the said theorem may informally be summarised as follows:

1. For elliptically distributed[5] portfolios, VaR is a coherent risk measure—it is monotonic, positive homogeneous and translation invariant, and it fulfils the sub-additivity property regardless of the probability level α used

2. In an elliptical world, among all portfolios with the same expected return, the portfolio minimizing VaR is the MV-minimizing portfolio. This statement applies to other risk measures that may or may not be coherent; in particular, it applies to the Lower Partial Moment even though this measure violets the positive homogeneity and translation invariance axioms.

3. For non-elliptically distributed portfolios, most (if not all) of the standard results from the previous statements do not hold; in particular, VaR is more than questionable

The second statement above appears to mitigate the common use of 'portfolio variance'—a non-coherent, non-downside measure of risk. But unfortunately, there is no guarantee that real portfolio returns are always going to be elliptically distributed. And in view of the compelling theory of coherency, one implication that this entails is clear: in the real-world, all normative models for portfolio optimization must, at the very least, have the potential for accommodating a coherent risk measure.

Closing Remarks. It is worth noting at the outset that not only does GENO have a real capacity for accommodating any risk measure, it also allows inclusion of extra criteria in the MV decision model that may be used to address other issues mentioned above; a brief summary of its capability profile is as follows:

▪ On Non-Investability of Allocations. The fact that GENO can handle discrete variables that may be binary, integer or real means that the 'non-investability' problem does not arise—this is discussed in §**4**.

▪ On Ad hoc Constraints. The GENO framework allows inclusion into the MV model of *ad hoc* constraints (such as the 5-10-40 German Investment Law) via binary variables or logic propositions—this is discussed in §**4**.

▪ On Instability of Solutions. Dealing with portfolio instabilities arising from estimation errors is within GENO's capacity. However, the emphasis in this paper shall be on a related and more practical issue, namely dealing with the "opportunity costs" of holding a particular portfolio in the face of estimation errors. To that end, a robust MV model is presented in §**5** and its qualities are tested in Example 4.

▪ On Error Maximisation. GENO allows augmentation of the standard MV criterion set to include functions that may be designed to minimise the adverse effects of estimation errors—this is discussed in §**5**.

▪ On Alternative Risk Measures. Evolutionary algorithms are very versatile: they can accommodate any objective function that can be mapped onto an ordinal evaluation function, and said objective function does not even have to be a closed-form expression [43, p.37]. Thus, although the numerical example presented in this paper adopts portfolio variance as the risk measure, the M-programs that may successfully be addressed by GENO may include various other types of measures that could be of the 'downside', 'non-downside' and / or 'coherent' or 'non-coherent' varieties, and including even those that can only be defined algorithmically.

---

[5] An elliptical distribution is one whose probability density is constant on ellipsoids—in the bivariate case, the contours lines of the density surface are ellipses. Elliptic distributions are symmetric and uni-modal, and they possess a useful property to portfolio theory, namely that a linear sum of elliptic distributions is also elliptic. Examples of elliptic distributions include the Gaussian, Laplace, Student's *t*, Cauchy and Logistic distributions.

3 Portfolio Optimization: Multi-objective Dynamic Models

Preamble. This section presents two alternative models of the portfolio allocation problem, both of which are multi-objective optimization problems defined over a discrete domain. The first is a single-period pseudo-dynamic model—the "pseudo-dynamics" arising from the fact that investment allocation decisions are made serially, but all at the start of the holding period; the second is a truly dynamic multi-period model in which it is assumed that investment decisions are made simultaneously at each decision stage. The second model may also be operated in 'single-period mode' by merely declaring the optimization horizon parameter as T = 1. The formulations assume that all asset returns are drawn from elliptical probability distributions, and that said distributions are also 'stationary'; see e.g., [27, p.23]. The following notation applies:

> **Notation**. The term 'Opt' denotes an operator whose operand is a discrete set of criterion functions; it is a command to 'optimize the operand'; it may be "distributed onto" the elements of the criterion set, or "factored out" from such a collection; when the operand is a singleton, then 'Opt' means 'minimize' or 'maximize' depending on the context. The associated term 'arg opt' means 'the argument that optimizes the operand'; when the operand is a singleton, it means 'arg min' or 'arg max' depending on the context.

The Sequential Decisions Model. Given an asset universe of size *n*, let the prime $'$ denote matrix transposition; let $\mathbf{u} = (u_1, u_2, \cdots, u_n)'$ denote the fractional sequentially applied investment allocations in the *n* assets; let $\mathbf{x} = (x_1, x_2, \cdots, x_{n+1})'$ represent the state of an investor's wealth, the component x_k denoting *the balance yet to be invested* as at "time" *k*; let the vector $\mathbf{r} = (r_1, r_2, \cdots, r_n)'$ denote the expected values for the *n* asset returns over a single holding period; and let symbol 'Opt' denote an operator that separately optimizes each element of a criterion set, say $\{f_1, f_2, \ldots, f_m\}$. If asset returns are random variables drawn from a joint probability distribution with covariance matrix $[c_{tk}]$, then the portfolio problem may be modelled as a bi-objective "dynamic" optimization problem on the discrete domain $\mathbf{N} = \{1, 2, \ldots, n\}$ as follows:

MP4:
$$\underset{\mathbf{u}}{\mathbf{Opt}}\left\{ g(\mathbf{u}),\, h(\mathbf{u}) \right\}$$

Subject to:
$$x_{k+1} = x_k - u_k$$
$$x_1 = 1$$
$$x_{n+1} = 0$$
$$\sum_{k=1}^{n} u_k = 1$$
$$u_k \in [0, 1]$$
$$k \in \{1, 2, \ldots, n\}$$

Where:
$$g(\mathbf{u}) = \langle \mathbf{r}, \mathbf{u} \rangle = \sum_{k=1}^{n} r_k u_k \quad \text{is the portfolio return;}$$

$$h(\mathbf{u}) = \sum_{t=1}^{n}\ \sum_{k=1}^{n} u_t u_k c_{tk} \quad \text{is the portfolio variance.}$$

Remarks. The M-program MP4 is a two-point boundary value problem, and the examples in [61] attest to GENO's efficacy on this class of problems. It is easy to implement when only standard constraints apply, as is the case above. But in the real-world, constraints pertaining to 'asset exposure', 'transaction costs' and 'investability' are often necessary, and these render MP4 rather awkward to implement. And it was this awkwardness that prompted the formulation of a second model explained next.

The Parallel Decisions Model. The 'parallel-decisions' model is a truly dynamic M-program in which the basic premise is that investment allocations are made in tandem at each decision stage; the formulation proceeds as follows. Given an asset universe of size n, let the matrix $\bar{\mathbf{u}}_k = [\mathbf{u}_0 \,|\, \mathbf{u}_1 \,|\cdots|\, \mathbf{u}_k]$ denote the fractional investment decisions made over time—the column vector $\mathbf{u}_k$ being applied at the k-th instance; let $\bar{\mathbf{x}}_k = [\mathbf{x}_I \,|\, \mathbf{x}_2 \,|\cdots|\, \mathbf{x}_k]$ denote a state matrix that describes the investor's asset holdings over time; let the column vectors of the matrix $\bar{\pi}_k = [\pi_1 \,|\, \pi_2 \,|\cdots|\, \pi_k]$ denote the expected single-period asset returns at time $t = 1, 2, \ldots, k$; assume that each vector π_k comprises n random returns r_k^i drawn from a joint probability distribution with covariance matrix $[c_k^{ij}]$, and that the said distributions is elliptical, independent and identical over time; let the prime $'$ denote matrix transposition; let h denote a multi-period measure of portfolio risk and g denote the portfolio's total return over the period $[0, T]$; then the multi-period portfolio problem may be modelled as a bi-objective dynamic optimization problem on the discrete time domain $k \in \{0, 1, 2, \ldots, T\}$ as:

MP5: $\qquad \underset{\bar{\mathbf{u}}}{\mathbf{Opt}} \left\{ g(\bar{\mathbf{u}}_T),\ h(\bar{\mathbf{u}}_T) \right\}$

$\qquad$ Subject to: $\quad x_{k+1}^j = x_k^j + u_k^j\,; \qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$

$$\sum_{j=1}^{n} u_k^j = 1\,; \qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$$

$$u_k^j \in [0, 1]\,; \qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$$

$$\mathbf{x}_0 = \mathbf{0}$$

$\qquad$ Where:[6] $\qquad g(\bar{\mathbf{u}}_T) = \mathrm{tr}\,[\bar{\pi}_{T+1}'\, \bar{\mathbf{u}}_T]\quad$ is the total portfolio return over $[0, T]$

$$h(\bar{\mathbf{u}}_T) \equiv \sum_{k=1}^{T} H_k^T(L_T, \tau)\ \text{ is a coherent measure of the portfolio risk over } [0, T].$$

Remarks. The M-program MP5 is a multi-period model, and as such one needs to specify a risk criterion that is 'time-consistent', i.e. one that ensures compatibility of consecutive decisions implied by the measure [37]. To that end, an appropriate measure is one formulated by Hardy & Wirch [26] called the 'Iterated Conditional Tail Expectation' (ICTE). The ICTE satisfies the coherency axioms of Artzner, et al. [3, 4], as well as the requirements of 'dynamic consistency' and 'relevancy' as articulated by Riedel [55]. The following is a description of the formula for the recommended risk measure h: assume the planning horizon $[0, T]$ is divided into n sub-intervals, each of length τ years; let elements of the set $\mathbb{N}_T \equiv \{1, 2, \cdots, T\}$ denote the number of iteration stages and define a 'stage locator' variable by $m = \eta_i \tau$; let Φ denote the cumulative distribution function of the standardized Gaussian random variable z; and let μ and σ be parameters of a log-normal asset price process that underlies a random sequence of loss variables L_k used to assess the portfolio's risk; then, at any decision stage $k \in [0, T] \bigcap \{m : m = \eta_i \tau,\ \eta_i \in \mathbb{N}_T\}$, we have that $k = T - m$, and the ICTE at confidence level α is (see **Appendix B**, Equation B11):

$$H_k^T(L_T, \tau) = L_k \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right]^{\left(\frac{m}{\tau}\right)} \exp\!\left(m\,[\mu + 0.5\sigma^2]\right) \tag{1}$$

The derivation of (1) together with a proposal on how it could be implemented is presented in **Appendix B**.

---

[6] The notation 'tr' denotes the matrix trace operator.

4 Handling Special Constraints

Preamble. Incorporating special constraints into the M-V model is the subject of much current research, but GENO affords some simple methods in this regard. Example 3 below—which involves 'exposure' and 'round-lot' constraints—is a case in point: in the GENO framework, such constraints are, in the main, implemented by merely entering the required numbers into the input data matrix. Otherwise, one can always include suitable *ad hoc* mathematical expressions for such special constraints as follows.

Modelling Methods. There are two other types of constraints that have garnered attention in the literature: (i) *cardinality constraints* specify upper and lower bounds on the size of the portfolio to be constructed from a given "asset universe", here denoted by the set **E**—such constraints may be required for portfolio diversity and efficient monitoring; (ii) *buy-in thresholds* specify the minimum amount that may invested in a particular asset—such constraints pertain to the need to minimise transaction costs. Both types are easily accommodated in the GENO framework: 'buy-in thresholds' are implemented by simply declaring a figure for the lower-bound variables, UCB and LSB (cf. Example 3) and cardinality constraints are dealt with as follows.

First note that, the mere act of declaring the size of the state vector establishes the size of the "asset universe" $|\mathbf{E}|$; if the asset composition of the proposed portfolio coincides with **E**, then cardinality constraints are irrelevant; otherwise, cardinality constraints may be implemented directly via binary variables, or using Boolean variables together with some logic propositions as follows.

The Direct Approach. To construct a portfolio of size $N \in [N_L, N_U]$, from the "asset universe" **E**, where N_L is minimum and N_U is the maximum portfolio size, one approach suggested by Stein, Branke & Schmeck [65] is to introduce into MP5 the following set of constraints in which L_k^j and U_k^j are the lower and upper bounds respectively for u_k^j:

$$u_k^j \in [b_k^j L_k^j, \ b_k^j U_k^j] \tag{2a}$$

$$\sum_{j=1}^{|E|} b_k^j = N \in [N_L, \ N_U] \tag{2b}$$

$$b_k^j \in \{0, 1\} \tag{2c}$$

Similarly, if **D** is the subset of the "asset universe" that is subject to the 5-10-40 rule under Germany Investment Law [22], then the '5-10-40 constraint' may be modelled as follows.

$$u_k^j - 0.05 b_k^j \leq 0.05 \qquad j \in \mathbf{D} \tag{3a}$$

$$\sum_{j \in \mathbf{D}} b_k^j u_k^j \leq 0.4 \tag{3b}$$

$$\sum_{j \in \mathbf{D}} b_k^j \leq |D| \tag{3c}$$

The GDP Approach. A much more expressive approach would be to adopt a modelling technique devised by Chemical Engineers at Carnegie-Mellon University called Generalised Disjunctive Programming (GDP) [25]. The method employs Boolean variables and logic propositions to express constraints of the 'either-or' type. For the sake of illustration, consider the set of relations:

$$b_k^j \quad \Rightarrow \quad u_k^j \in [L_k^j, \ U_k^j] \tag{4a}$$

$$\neg b_k^j \quad \Rightarrow \quad u_k^j \in \{0\} \tag{4b}$$

$$b_k^1 \vee b_k^2 \cdots \vee b_k^{|E|} = \text{True} \tag{4c}$$

$$f(\mathbf{b}) = \text{True} \tag{4d}$$

$$b_i \in \{\text{True, False}\} \tag{4e}$$

The Boolean variable b_k^j indicates whether the *j*-th asset is *in* (b_k^j = True), or *out* (b_k^j = False) of a particular portfolio; equation (4c) ensures that the proposed portfolio contains at least one asset; and the function *f* (**b**) is a proposition that stipulates further conditions; for example, *f* might contain terms such as $b_k^m \veebar b_k^n$ (in which $\veebar$ is the 'exclusive-OR' operator) to implement a diversity requirement that states: "assets *m* and *n* cannot simultaneously be part of any particular portfolio"; and of course, several such propositions are allowed.

Closing Remarks. Note that both methods outlined above ultimately result in a 'Mixed-Binary Program', and the numerical examples in [61] show that GENO is well capable of solving this class of problems.

5 Accommodating Uncertainty

Preamble. The solution to a portfolio problem is not computable unless or until one knows the model parameters, i.e. the mean vector and the covariance matrix of the asset returns. In practice therefore, portfolio optimization is usually a two-step procedure that proceeds as follows: given some time series data on asset returns covering a historical period, (i) statistical estimates of the mean vector and covariance matrix of the asset returns are computed based on the observed data; (ii) the estimates computed at Step 1 are then regarded as if they were the true population parameters for the *unknown* probability distribution that generated the observed data; they are plugged into the MV model, and the latter is solved to produce the (supposedly) optimal asset allocations. A potential source of weakness in this plug-in solution strategy is at Step 1, the statistical estimation stage.

In practice, unless there are strong *a priori* reasons to suggest otherwise,

> ". . . it is usually assumed that the observations are [Gaussian] distributed. This assumption is partly justified by an appeal to the Central Limit Theorem. However, [the Gaussian distribution] can rarely be taken for granted, and the famous remark of Poincaré remains apt: '. . . everyone believes in the [Gaussian] laws of errors, the experimenters because they think it is a mathematical theorem, the mathematicians because they think it is an empirical fact' " [Paraphrased from 27, p.116]

But once the Gaussian assumption is cast into doubt, then robustness of the estimators employed becomes a key issue: it turns out that some estimators suffer a drastic loss of statistical efficiency[7] if the data supplied them is not exactly Gaussian. This is mainly what lies at the root of the 'instability-of-optimized-portfolios' problem identified by Michaud [46]—a re-calibration of an MV solver using a non-robust estimator and new data that is "not as Gaussian" as the original sample results in significantly different estimates for the model parameters, and hence solver generates asset allocations that are quiet removed from the old ones.

---

[7] A good statistical estimator is one that is unbiased and has a small mean square error; if an estimator is unbiased, then its mean square error coincides with its variance, and so given two estimators that are unbiased, a choice of the one with the lower variance is consistent with minimising the mean square error. This goodness criteria is measured by 'relative efficiency': an estimator $\mathbf{e}_1$ of a parameter $\mathbf{p}$ is said to be relatively more efficient than another estimator $\mathbf{e}_2$ if both $\mathbf{e}_1$ and $\mathbf{e}_2$ are unbiased estimators of $\mathbf{p}$, but the variance of $\mathbf{e}_1$ is less than that of $\mathbf{e}_2$.

There are several ways of tackling this problem, each with its own merits and demerits; however a detailed discussion of these methods is beyond the scope of this paper; interested readers may wish to consult a recent survey in [20]. The most obvious approach is to use robust estimators at Step 1 of the "plug-in" strategy; a second approach is to formulate robust version of the MV problem assuming certain characteristics for the asset return distributions; a third approach essentially combines the two stages of the plug-in strategy—portfolios are generated using optimal estimation in one single step such as in [12]; a fourth approach is to combine the investor's prior beliefs on asset returns with evidence from historical observations and then use estimators based on Bayes' theorem.

Modelling Method. The method proposed in this paper is essentially a plug-in strategy in the vein of Tütüncü & Koenig [70]: its basic rationale is that, since statistical point estimates are *always* accompanied by some degree of imprecision *no matter how sophisticated the estimator used* may be, it is prudent to use intervals or 'uncertainty sets' within which the parameters in question may be assumed or shown to lie with a certain degree of confidence. The method is an adaptation of a nonlinear interval programming (NLIP) technique for dealing with data uncertainty that is presented by Jiang, Han, Liu & Liu [32]. The formulation is as follows.

Consider a generic M-program in which the constraint and objective functions are determined by a decision vector $\mathbf{u} \in \mathbf{U} \subseteq \mathbb{R}^n$ conditional on a data vector $\mathbf{d} \in \Xi \subseteq \mathbb{R}^m$, viz.:

MP6a: $$\underset{\mathbf{u}}{\text{Max}} \left\{ f_0(\mathbf{d}, \mathbf{u}) \,\middle|\, f_1(\mathbf{d}, \mathbf{u}) \leq 0; \ \mathbf{d} \in \Xi, \mathbf{u} \in \mathbf{U} \right\}$$

In practice, there is always some degree of uncertainty as to the true value of the vector $\mathbf{d}$. This ignorance may be modelled by setting the elements of $\mathbf{d}$ as closed intervals $\mathrm{id}_q \in \mathbb{IR}$, $q = \{1, 2, \ldots, m\}$ on $\mathbb{R}$ rather than single points. And the introduction of the intervals id_q into MP6a implies that the functions f and c also assume interval-like qualities, and one may characterise them by the mid-points and radii, or by the end-points on their respective interval ranges. Exactly how the intervals id_q propagate through the functions f_0 and f_1 may be worked out analytically or empirically using the calculus of (real) interval numbers [36, 48].

Jiang, et al. [32] base their NLIP formulation on a preference relation for closed intervals in $\mathbb{IR}$ that was first suggested by Ishibuchi & Tanaka [31]; a simplified definition of the said relation is as follows:

- DEFINITION 1 [Interval Preference Relation] Consider two closed intervals $\mathbf{A} = [a_L, a_R]$ and $\mathbf{B} = [b_L, b_R]$ on the real line that pertain to the values attained by a criterion function of an optimization problem. The intervals may also be characterised by an ordered pair comprising the centre or mid-point $m_a \equiv 0.5(a_L + a_R)$, and the radius or half-width $w_a \equiv 0.5(a_R - a_L)$, viz., $\mathbf{A} = (m_a, w_a)$ and $\mathbf{B} = (m_b, w_b)$; a useful identity is: $[a_L, a_R] \equiv m_a + w_a [-1, 1]$. If the associated optimization problem is one of maximization, then the interval preference relation $\mathbf{A} \succ_P \mathbf{B}$ (read: '$\mathbf{A}$ is better than $\mathbf{B}$') implies the following:

$$\mathbf{A} \succ_P \mathbf{B} \iff (m_a \geq m_b) \wedge (w_a \leq w_b) \tag{5a}$$

  Similarly, if the associated optimization problem is one of minimization, then $\mathbf{A} \succ_P \mathbf{B}$ implies the following:

$$\mathbf{A} \succ_P \mathbf{B} \iff (m_a \leq m_b) \wedge (w_a \leq w_b) \tag{5b}$$

- REMARKS 1   This is a simplified version of the preference relation in [31] that retains only the essential elements.

Let f_{0L} and f_{0R} denote the left and right end-points, respectively, of the objective function interval induced by the introduction of the data intervals $\mathrm{id}_q \in \mathbb{IR}$, $q = \{1, 2, \ldots, m\}$; let f_{1L} and f_{1R} be similarly defined for the constraint function f_1; and let m_0 and w_0 denote the mid-point and radius functions with respect to the intervals of $f_0(\mathbf{u})$; then clearly:

$$f_{0L}(\mathbf{u}) \equiv \underset{\mathbf{d}}{\mathrm{Inf}}\{f_0(\mathbf{d},\mathbf{u})\} \tag{6a}$$

$$f_{0R}(\mathbf{u}) \equiv \underset{\mathbf{d}}{\mathrm{Sup}}\{f_0(\mathbf{d},\mathbf{u})\} \tag{6b}$$

Note that $f_{0L}(\mathbf{u})$ and $f_{0R}(\mathbf{u})$ are deterministic because the uncertainty associated with the data vector $\mathbf{d}$ is effectively eliminated by the 'Inf' and 'Sup' operations. And by definition:

$$m_0(\mathbf{u}) \equiv 0.5(f_{0L}(\mathbf{u}) + f_{0R}(\mathbf{u})) \tag{7a}$$

$$w_0(\mathbf{u}) \equiv 0.5(f_{0R}(\mathbf{u}) - f_{0L}(\mathbf{u})) \tag{7b}$$

The preference relations in DEFINITION 1 implicitly comprises two choice measures: (i) a cardinal preference encapsulated by the order relation on the interval mid-points; (ii) an aversion for uncertainty encapsulated by the order relation on the interval widths. Accordingly, if one applies the preference relation in DEFINITION 1 to the functions f_0 and f_1 in MP6a, then one obtains a robust version of the original M-program, namely the interval-induced, bi-objective optimization problem that may be stated thus:

MP6b: $$\underset{\mathbf{u}}{\mathrm{Opt}}\left\{m_0(\mathbf{u}), w_0(\mathbf{u}) \,\middle|\, f_{1R}(\mathbf{u}) \le 0;\ \mathbf{d} \in \mathbb{IR}^{\,m};\, \mathbf{u} \in \mathbf{U}\right\}$$

The mid-point function $m_0(\mathbf{u})$ is correlated to the expected value of the partial function $f_0(\mathbf{u})$, and the radius function $w_0(\mathbf{u})$ is correlated to the variance $\mathbb{V}[\,f_0\,]$. For example, if one assumes, purely for the sake of argument, that f_0 is uniformly distributed on $[f_{0L}, f_{0R}]$, then:

$$m_0(\mathbf{u}) \equiv \mathbb{E}[\,f_0(\mathbf{u})\,] \tag{8a}$$

$$w_0(\mathbf{u}) \equiv \sqrt{3\mathbb{V}\big[f_0(\mathbf{u})\big]} \tag{8b}$$

In the case of MP6b therefore, the appropriate interpretation of the operator 'Opt'—when "distributed" over the criterion set $\{\,m_0, w_0\,\}$ in MP6b is that it assumes the meanings 'maximize' and 'minimize', respectively because maximizing m_0 is consistent with the original program MP6a, and minimizing w_0 is consistent with reducing the adverse effects of the uncertainty in $\mathbf{d}$. Thus, MP6b is in effect a reward-risk optimization formulation—maximizing the mid-point function m_0 also maximizes the expected value of f_0, i.e. 'the reward'; and insofar 'variance' measures risk, then minimizing the radius function w_0 builds into the solution a certain degree of robustness to the estimation risk emanating from variations in the data vector $\mathbf{d}$. This fact may be used to formulate a robust version of the portfolio problem. For example, consider a single-period version of MP5: the uncertain data are the asset returns r_i and each element of the covariance matrix $[c_{ij}]$; an application of DEFINITION 1 separately to the return function g and the risk function h would result in a criterion set comprising four elements; the robust version of MP5 may be stated thus:

MP7: $$\underset{\mathbf{u}}{\mathrm{Opt}}\left\{m_g(\overline{\mathbf{u}}),\ w_g(\overline{\mathbf{u}}),\ m_h(\overline{\mathbf{u}}),\ w_h(\overline{\mathbf{u}})\right\}$$

Subject to: $\quad x_{k+1}^{j} = x_k^{j} + u_k^{j}\,; \qquad\qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$

$$\sum_{j=1}^{n} u_k^{j} = 1\,; \qquad\qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$$

$\qquad\qquad\quad u_k^{j} \in [0,1]\,; \qquad\qquad\quad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\}$

$\qquad\qquad\quad \mathbf{x}_0 = \mathbf{0}$

Where at each decision stage k:

$$m_g(\mathbf{u}) \equiv 0.5(g_L(\mathbf{u}) + g_R(\mathbf{u}));$$

$$w_g(\mathbf{u}) \equiv 0.5(g_R(\mathbf{u}) - g_L(\mathbf{u}));$$

$$m_h(\mathbf{u}) \equiv 0.5(h_L(\mathbf{u}) + h_R(\mathbf{u}));$$

$$w_h(\mathbf{u}) \equiv 0.5(h_R(\mathbf{u}) - h_L(\mathbf{u})).$$

MP7 may seem rather complicated to implement due to the fact that the requisite operands, namely the 'end-point' functions $g_L(\mathbf{u})$, $g_R(\mathbf{u})$, $h_L(\mathbf{u})$ and $h_R(\mathbf{u})$, are defined via optimization processes (according to equation 6). But this problem is more apparent than real: the data components feature linearly in both g and h, and so it is easy to derive explicit expressions for the 'end-point' functions using interval analyses of the functions in question. Note that for the single-holding period version of MP5, the portfolio return function "degenerates" to the inner product $g = \langle \mathbf{u}, \mathbf{r} \rangle$ which is bilinear. And so if the vectors $\mathbf{r}_L$ and $\mathbf{r}_R$ denote collections of the left and right-end points of the asset returns vector $\mathbf{r}$, then from the natural interval extension of g with respect to $\mathbf{r}$, it follows immediately that: [8]

$$g_L(\mathbf{u}) \equiv \langle \mathbf{u}, \mathbf{r}_L \rangle \tag{9a}$$

$$g_R(\mathbf{u}) \equiv \langle \mathbf{u}, \mathbf{r}_R \rangle \tag{9b}$$

One does not need to explicitly determine the 'end-point' pertaining to h because the functions ultimately required by MP7—i.e. m_h and w_h—may be stated directly using the definitions of 'matrix mid-point' and 'matrix width'; the former is exactly the definition in (7a) applied to each interval element of a matrix; the latter is a scalar defined as the supremum norm of a matrix; see [48, §7.1]. To that end, let c_{ijL} and c_{ijR} denote the left and right end-points of the interval pertaining to the covariance matrix element, c_{ij} and let the w_c denote the width of the covariance matrix $[c_{ij}]$; then it is easy to show that:

$$m_h(\mathbf{u}) \equiv \mathbf{u}'\,[0.5(c_{ijL} + c_{ijR})]\,\mathbf{u} \tag{10a}$$

$$w_h(\mathbf{u}) \equiv w_c\langle \mathbf{u}, \mathbf{u} \rangle \tag{10b}$$

The width function w_c, which as mentioned earlier is defined as the supremum norm of a matrix, is:

$$w_c \equiv \underset{i,j}{\mathrm{Max}}\left\{ 0.5(c_{ijR} - c_{ijL}) \right\} \tag{10c}$$

Closing Remarks. Note that if the interval vector $\mathbf{d}$ is degenerate—i.e. when all its elements are specific constants and there is no uncertainty about their actual values—then all width functions vanish; equations (9a) and (9b) coalesce into the normal portfolio return $g = \langle \mathbf{u}, \mathbf{r} \rangle$; equation (10a) becomes the normal portfolio variance; and thus MP7 reverts to the normal, non-robust formulation in MP5. Note also that m_h and w_h are just as easily derived for the coherent risk function recommended for MP5 because the data elements, μ and σ, are arguments of monotonic functions; see [48, §5.2]. And in all cases, the correct interpretation of 'Opt' when "distributed" over the criterion set $\{ m_g, w_g, m_h, w_h \}$ is 'max', 'min', 'min' and 'min' respectively.

---

[8] See [48, Chap. 5]. An alternative approach would be to use a 'calculus-of-variations' argument as follows:

$$g = \langle \mathbf{u}, \mathbf{r} \rangle$$

$$\delta g = \langle \mathbf{u}, \delta \mathbf{r} \rangle$$

$$g_R - g_L = \langle \mathbf{u}, (\mathbf{r}_R - \mathbf{r}_L) \rangle$$

$$g_R - g_L = \langle \mathbf{u}, \mathbf{r}_R \rangle - \langle \mathbf{u}, \mathbf{r}_L \rangle \;\Leftrightarrow\; (g_R(\mathbf{u}) \equiv \langle \mathbf{u}, \mathbf{r}_R \rangle) \wedge (g_L(\mathbf{u}) \equiv \langle \mathbf{u}, \mathbf{r}_L \rangle)$$

6 Solution of the Models

Preamble. The notion of optimization is unambiguous in the uni-objective context: the verb 'optimize' is a command that is universally understood to mean 'compute the extrema of some criterion function'. But the same cannot be said of multi-objective problems because the various criterion functions involved evaluate candidate solutions in disparate ways: taken pair-wise, some may "compete" in the sense that an improvement with respect to one degrades the solution as assessed the other; others may "collude" in the sense that an improvement in one entails the same in the other; and others may be totally independent. It is not surprising therefore that there exists different types of optima for such problems and furthermore, some of the optimization notions yield non-unique solutions. A common approach to the quandary posed by the non-uniqueness problem is to introduce a *model-user* or *decision-maker*—i.e. the person for whom the computed solutions are intended—and require such a person to articulate preferences at some stage during the solution process. But this paper shall only be concerned with one—the 'compromise solution'. The rationale for the compromise solution is presented in [62] and shall not be repeated here.

The Compromise Solution and its Computation. The compromise solution concept that was first introduced by Salukvadze [58] and later independently presented by Yu [76] and Zeleny [77]. It is based on the common-sense and compelling notion that the best option is a feasible point that yields values that are closest to an ideal outcome—the *ideal* being that point at which each criterion is optimized to the fullest extent possible. The rationale for the compromise solution is best explained in terms of the two-dimensional outcome space depicted in Figure 1 below in which $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are finite-valued criterion functions of a decision vector $\mathbf{x}$, and the collection of all such feasible outcomes constitutes the outcome set $\mathbf{\Omega}$.

Associated with each outcome vector $\mathbf{\omega}$ are four 'boundary' outcome vectors $\mathbf{\omega}_1$, $\mathbf{\omega}_2$, $\mathbf{\omega}_3$ and $\mathbf{\omega}_4$; these are points where lines that are parallel to the axes f_1 and f_2 intersect the boundary of the set $\mathbf{\Omega}$, which shall hereafter be denoted by $\partial\mathbf{\Omega}$. The vertices of the smallest rectangle enclosing $\mathbf{\Omega}$ comprise the 'utopia set', and for any given vertex vector $\mathbf{z}_n$, each of its dimensions represents the best possible outcome, i.e. the *global* solution, that could be attained by maximizing or minimizing a particular criterion independently. However, only one vertex would be relevant in any given scenario and such a vertex is conventionally called the 'ideal point'. In Figure 1 below, $\mathbf{z}_1$ is the ideal point when both criteria are required to be minimized; whereas $\mathbf{z}_4$ is for the case where criterion 1 is to be minimized, and criterion 2 maximized.

Figure 1: Outcome Set of a bi-objective Optimization Problem
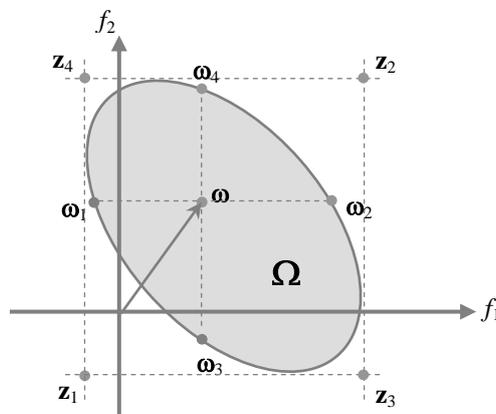
One may define the compromise solution in two stages as follows:

▪ DEFINITION 2 [The Ideal Point]:   Let $\mathbf{\Omega} \subset \mathbf{R}^n$ denote an outcome set; let L_j and U_j denote the lower and upper bounds respectively for the criterion f_j at $\mathbf{\omega}$ assuming all other outcomes remain constant; let $\mathbf{z}_i$, denote the *ideal point* for the problem at hand, then the coordinates of $\mathbf{z}_i$ are given by scalars z_{ij} defined as:

$$z_{ij} = \begin{cases} \underset{\omega \in \Omega}{\text{Sup}} \left\{ U_j(\mathbf{\omega}) \right\}, & \text{if the } j\text{th criterion requires maximizing} \\ \\ \underset{\omega \in \Omega}{\text{Inf}} \left\{ L_j(\mathbf{\omega}) \right\}, & \text{if the } j\text{th criterion requires minimizing} \end{cases} \qquad (11a)$$

REMARKS 2: Because ideal outcomes are normally not *jointly* attainable, a compromise is required.

▪ DEFINITION 3 [The Compromise Solution]:   Let the point $\mathbf{z}_i$ be the *ideal point* a given multi-objective optimization problem; then the compromise solution is a member of those feasible controls whose outcomes are closest to the ideal outcome as measured by some distance function such as the Euclidean or Tchebycheff metric; thus, in terms of the Tchebycheff metric, the compromise solution is a feasible vector $\mathbf{x}^*$ whose corresponding outcome vector $\mathbf{\omega}^*$ belongs to a set of outcomes $\xi \subset \partial\mathbf{\Omega}$ that is defined as:

$$\xi(\mathbf{z}_i) = \left\{ \mathbf{\omega} \in \partial\mathbf{\Omega} : \mathbf{\omega} = \arg\min ||\mathbf{\omega} - \mathbf{z}_i||_\infty \right\} \qquad (11b)$$

REMARKS 3: The definition of $\xi(\mathbf{z}_i)$ entails two processes: (i) the obvious minimization process denoted by the 'arg min' operator; (ii) the less obvious search process that is supposed to delineate the boundary set $\partial\mathbf{\Omega}$. The rationale underlying this solution concept may be explained as follows: (a) there is no question that, if it were achievable, the ideal outcome vector $\mathbf{z}_i$ would constitute *the* optimal solution to the multi-objective optimization problem under study; (b) but since this is usually not the case, one has to "compromise downwards" from the ideal outcome $\mathbf{z}_i$ to a less-than-ideal outcome $\mathbf{\omega}^*$ that corresponds to a feasible vector $\mathbf{x}^*$ and obviously, the extent of the "downward compromise", i.e. the quantity of criterion value that must be given up along each dimension, has to be minimal, hence the distance-minimizing operation in the definition of the solution set ξ, and the stipulation that $\mathbf{\omega}^* \in \xi$.

Closing Remarks. The compromise solution is logically sound and compelling; it dispenses with the need to calibrate utility functions as is the case in other single-point MV solution techniques such as [15]; the model-user is not even required to express preferences at any point in the solution process; rather, implicit to the method is an assumption of investor behaviour that effectively asserts the following:

Assumption: The model-user is *rational* and would naturally prefer the ideal solution; but since such a solution is usually not attainable, the model-user *will* accept the compromise solution as being the best that can be done in the given circumstances.

In addition, unlike many user-mediated multi-objective solution methods that

"tend to become cumbersome and even useless as we increase the number of objectives" [10, p.4]

the size of the problem, as measured by the dimension of $\mathbf{\Omega}$, is largely irrelevant.

In order to implement the compromise solution, one must first evaluate (11a) and this can be a serious huddle in practice. If it is possible to reformulate the criterion functions f_i such that $\mathbf{z}_i \equiv \mathbf{0}$, then that is what should be done; failing that, then a technique involving a monotonic function called 'The Generalized Loss Transform' (or *g*-loss transform) may be used instead—the *g*-loss transform is explained in detail [62]. But assuming the ideal vector $\mathbf{z}_i$ is known, then the 'arg min' operation in (11b) may be viewed as a command to find a solution to a nonlinear equation system, and so the methods explained in [63] are applicable; the two uni-objective methods presented there, namely the composite-metric method (code name: EDR) and the NCP method (code name: ECR), are used on Example 4, primarily for the sake of efficiency.

7 Numerical Examples

Preamble. This section presents some numerical examples whose purpose is to assess the ideas and techniques presented in this paper. The requisite input data is from various sources: the first example is from [40] and it provides a basic comparison of the solution method employed there with GENO; the second example is from [44] and it serves to compare GENO to the Black-Litterman model as well as the 're-sampled efficiency' method of Richard and Robert Michaud;[9] the third example is from [19]—it illustrates the handling of non-traditional *ad hoc* constraints and assesses GENO's solution *per se*; the last example employs data from [70] and serves to assess the robust optimization model formulated and presented in §**5** above.

Example 1: A single-period MV optimization problem [10]

MP8: $\mathbf{Opt}_{\mathbf{u}} \left\{ g(\mathbf{u}), h(\mathbf{u}) \right\}$

Subject to: $x_{k+1} = x_k - u_k; \quad k \in \{1, 2, \dots, 5\}$

$x_1 = 1$

$x_{n+1} = 0$

$\sum_{k=1}^{n} u_k = 1$

$u_k \in [0, 1]$

Where: $g(\mathbf{u}) = \langle \mathbf{r}, \mathbf{u} \rangle = \sum_{k=1}^{5} r_k u_k$ is the portfolio return;

$h(\mathbf{u}) = \sum_{t=1}^{5} \sum_{k=1}^{5} u_t u_k c_{tk}$ is the portfolio variance.

Tables 1A: Expected Annual Returns (%) [40, p.163]

Asset Name	SECURITY 1	SECURITY 2	SECURITY 3	SECURITY 4	SECURITY 5
MEAN RETURN	15.1	12.5	14.7	9.02	17.68

Tables 1B: Covariance Matrix [40, p.163]

Covariance Matrix	SECURITY 1	SECURITY 2	SECURITY 3	SECURITY 4	SECURITY 5
SECURITY 1	2.30	0.93	0.62	0.74	-0.23
SECURITY 2	0.93	1.40	0.22	0.56	0.26
SECURITY 3	0.62	0.22	1.80	0.78	-0.27
SECURITY 4	0.74	0.56	0.78	3.40	-0.56
SECURITY 5	-0.23	0.26	-0.27	-0.56	2.60

---

[9] US Patent #6,003,018, December 1999. A brief description of the re-sampled efficiency method may be found in Becker, Gürtler & Hibbeln [6] who present an extensive simulation study that compares it to the traditional Markowitz approach.

[10] MP6 is a sequential-decisions model; this was used to generate the first set of result reported below; the second set of results was generated by a parallel-decisions model which, for the sake brevity, is *not* re-stated here.

GENO **Output** [The Sequential Decisions Model]

Generation Number	Time (sec)	Portfolio "Loss" [11]	Portfolio Risk
0	0.00	0.06452204	0.95838589
10	4.13	0.06935328	0.79053227
20	4.10	0.06935326	0.79053228
30	4.15	0.06935326	0.79053228
40	4.15	0.06935326	0.79053228
50	4.13	0.06935326	0.79053228
60	4.09	0.06935326	0.79053228
70	4.12	0.06935326	0.79053228
80	4.09	0.06935326	0.79053228
90	4.12	0.06935326	0.79053228
100	4.07	0.06935326	0.79053228

Asset Name:	SECURITY 1	SECURITY 2	SECURITY 3	SECURITY 4	SECURITY 5
Allocation:	0.088790530	0.250125980	0.282629880	0.103244140	0.275209470

Portfolio Return (%): 14.41893378
Portfolio Risk: 0.79053228
Return-Risk Ratio: 18.23952574

GENO **Output** [The Parallel Decisions Model]

Generation Number	Time (sec)	Portfolio "Loss"	Portfolio Risk
0	0.00	0.06748238	0.80007756
10	4.70	0.06745308	0.79942382
20	4.73	0.06745308	0.79942382
30	4.71	0.06745308	0.79942382
40	4.79	0.06745308	0.79942382
50	4.87	0.06745308	0.79942382
60	4.90	0.06745308	0.79942382
70	4.76	0.06745308	0.79942382
80	4.85	0.06745308	0.79942382
90	4.84	0.06745308	0.79942382
100	4.67	0.06745308	0.79942382

Asset Name:	SECURITY 1	SECURITY 2	SECURITY 3	SECURITY 4	SECURITY 5
Allocation:	0.174000000	0.171999970	0.273000000	0.081000030	0.300000000

Portfolio Return (%): 14.82511990
Portfolio Risk: 0.79942382
Return-Risk Ratio: 18.54475636

Remarks on Example 1. Luenberger initially solves this problem by quadratic programming, i.e. using MP3a, the minimum variance version of the MV problem; he then generates a second efficient portfolio using the two-fund theorem [40, p.163]. As can be seen above, both GENO models generate sequences that converge, but to two different portfolios; the four portfolios are analysed and compared in Table 9 of §**8**.

---

[11] The Portfolio "Loss" is here defined as the reciprocal of the total portfolio return. This is so that the origin of the outcome space becomes the 'ideal point' which is required in the compromise solution scheme (see §**6**). This is permissible when portfolio return are non-negative (as is the case for all the examples in this paper); but for a more general approach, see 'The generalized Loss Transform' presented in [62].

Example 2: A single-period MV optimization problem

MP9:

$$\mathbf{Opt}_{\mathbf{u}} \left\{ g(\mathbf{u}), h(\mathbf{u}) \right\}$$

Subject to: $\quad x_{k+1} = x_k - u_k$

$$x_1 = 1$$

$$x_{n+1} = 0$$

$$\sum_{k=1}^{n} u_k = 1$$

$$u_k \in [0, 1]$$

$$k \in \{1, 2, \ldots, 8\}$$

Where: $\quad g(\mathbf{u}) = \langle \mathbf{r}, \mathbf{u} \rangle = \sum_{k=1}^{8} r_k u_k \quad$ is the portfolio return;

$$h(\mathbf{u}) = \sum_{t=1}^{8} \sum_{k=1}^{8} u_t u_k c_{tk} \quad \text{is the portfolio variance.}$$

Tables 2A: Expected Returns and Standard Deviations (%) [44]

Asset Class	Euro Bonds	US Bonds	Canada	France	Germany	Japan	UK	US
Mean Return	3.2	3.0	4.6	10.5	6.4	10.5	9.5	8.5
Std Deviation	5.4	7.0	19.0	24.4	21.5	24.4	20.8	14.9

Tables 2B: Correlation Matrix [44]

Correlation	Euro Bonds	US Bonds	Canada	France	Germany	Japan	UK	US
Euro Bonds	1.00	0.92	0.33	0.26	0.28	0.16	0.29	0.42
US Bonds	0.92	1.00	0.26	0.22	0.27	0.14	0.25	0.36
Canada	0.33	0.26	1.00	0.41	0.30	0.25	0.58	0.71
France	0.26	0.22	0.41	1.00	0.62	0.42	0.54	0.44
Germany	0.28	0.27	0.30	0.62	1.00	0.35	0.48	0.34
Japan	0.16	0.14	0.25	0.42	0.35	1.00	0.40	0.22
UK	0.29	0.25	0.58	0.54	0.48	0.40	1.00	0.56
US	0.42	0.36	0.71	0.44	0.34	0.22	0.56	1.00

Notes on Example 2. Michaud, et al. [44] use this example to assess their 're-sampled efficiency' method of portfolio optimization against the Black-Letterman model; in Table 2C below, the GENO results are appended to those from [44] for a three-way comparison; further comments may be found under Table 6 in §**8**.

GENO **Output**

Generation Number	Time (sec)	Portfolio "Loss"	Portfolio Risk
0	0.00	0.15815278	0.11624262
10	9.83	0.12826464	0.11276678
20	9.95	0.12826551	0.11276579
30	9.86	0.12826551	0.11276579
40	9.72	0.12826551	0.11276579
50	9.72	0.12826551	0.11276579
60	9.73	0.12826551	0.11276579
70	9.64	0.12826551	0.11276579
80	10.05	0.12826551	0.11276579
90	10.25	0.12826551	0.11276579
100	9.95	0.12826551	0.11276579

Asset Name:	Euro Bonds	US Bonds	Canada	France	Germany	Japan	UK	US
Allocation:	0.24238574	0.0000000	0.000000	0.07460059	0.000000	0.19568359	0.04040430	0.44692578

Portfolio Return (%): 7.79632824
Portfolio Risk: 0.11276579

Tables 2C: Comparative performance of portfolios: GENO versus others from [44, Table 3]

	Market (%)	Black-Litterman (%)	Michaud MSR (%)	GENO (%)
Euro Bonds	20.0	20.0	23.0	24.2
US Bonds	20.0	20.0	19.9	0.00
Canada	6.00	6.00	9.90	0.00
France	6.00	0.00	4.30	7.50
Germany	6.00	1.50	4.70	0.00
Japan	6.00	6.00	6.60	19.6
UK	6.00	1.50	5.40	4.00
US	30.0	45.0	26.2	27.5
PORTFOLIO RETURN (%)	6.10	5.40	5.90	7.80
PORTFOLIO RISK (%)	9.60	9.50	9.30	11.3
RETURN-RISK RATIO	0.64	0.57	0.63	0.69

Remarks on Example 2. In their evaluation of their re-sampling method against the Black-Litterman model, one conclusion that Michaud, Escher & Michaud draw is that

"under the same conditions, the Michaud MSR portfolio is better diversified, less benchmark centric, and less subject to large risky allocations" [Paraphrased from 44, p.12]

Admittedly, the GENO portfolio is also less diversified than the Michaud MSR. But this apparent weakness is more than compensated for by the superior return-risk ratio;[12] and in any case, the GENO framework allows one to easily induce more diversity via exposure constraints as illustrated by Example 3 below.

---

[12] Roy [57] was the first to suggest a variant of the reward-risk ratio called the 'Safety-First Ratio' as a measure for evaluating an investment strategy's value; Sharpe later applied Roy's ideas to the mean-variance framework of Markowitz and developed a metric now commonly known as the 'Sharpe Ratio', one of the best known performance evaluation measures in portfolio management [60]; the reward-risk ratio is equivalent to the Sharpe ratio when the benchmark return associated with the latter is zero, such as is the case when the benchmark asset is 'Cash' [14]; a portfolio with a relatively high reward-risk ratio awards larger returns for each additional unit of risk [19].

Example 3: A single-period MV optimization problem with exposure and round-lot constraints

MP10: $\underset{\overline{\mathbf{u}}}{\mathbf{Opt}} \left\{ g(\overline{\mathbf{u}}_{T\text{-}1}),\ h(\overline{\mathbf{u}}_{T\text{-}1}) \right\}$

Subject to:
$$x_{k+1}^{j} = x_{k}^{j} + u_{k}^{j}\,; \qquad\qquad j \in \{1, 2, 3, 4\}; \qquad k \in \{0\}$$

$$\sum_{j=1}^{n} u_{k}^{j} = 1\,; \qquad\qquad j \in \{1, 2, 3, 4\}; \qquad k \in \{0\}$$

$$u_{k}^{j} \in [0, 1]\,; \qquad\qquad j \in \{1, 2, 3, 4\}; \qquad k \in \{0\}$$

$$u_{k}^{j} \in [0, 1] \cap \{u_{k}^{j} : u_{k}^{j} = 0.13n,\ n \in \mathbb{Z}\}\,; \qquad j \in \{3\}; \qquad k \in \{0\}$$

$$\mathbf{x}_0 = \mathbf{0}\,;$$

$$T = 1\,;$$

$$n = 4.$$

Where:
$$g(\overline{\mathbf{u}}_{T\text{-}1}) = \mathrm{tr}\,[\overline{\pi}_{T}'\,\overline{\mathbf{u}}_{T\text{-}1}] \qquad \text{is the portfolio return;}$$

$$h(\overline{\mathbf{u}}_{T\text{-}1}) = \sum_{i=1}^{n}\ \sum_{j=1}^{n} c_{k+1}^{ij} u_{k}^{i} u_{k}^{j} \qquad \text{is the portfolio variance.}$$

Tables 3A: Expected Returns and Standard Deviations [19]

Asset Class	US Bonds	US Large-caps Equity	US Small-caps Equity	EAFE International Equity
Mean Rate of Return (%)	6.4	10.8	11.9	11.5
Standard Deviation	4.7	14.9	19.6	17.2

Tables 3B: Correlation Matrix [19]

Correlation Coefficients	US Bonds	US Large-caps Equity	US Small-caps Equity	EAFE International Equity
US Bonds	1.00	0.32	0.06	0.17
US Large-caps Equity	0.32	1.00	0.76	0.44
US Small-caps Equity	0.06	0.76	1.00	0.38
EAFE International Equity	0.17	0.44	0.38	1.00

Notes on Example 3. This example is a single-period implementation of the 'parallel-decisions' model that serves to illustrate, *inter alia*, the ease with which one may implement 'exposure' and 'round-lot' constraints in the GENO framework. Exposure constraints typically places a subjective upper limit on the investment proportion allocated to specific assets, and round-lot constraints are attempts to comply with discreteness requirements for investments in certain assets. In this particular case, there is an upper bound or exposure constraint on one asset denoted by u_1 and a round-lot constraint on a second asset denoted by u_3. The first constraint is implemented via data supplied to GENO (see Table 3C & 3D), whereas the second is implemented by a combination of an explicit *discrete* set-constraint (constraint #4 in MP10) as well as in the data supplied.

GENO **Output**

Generation Number	Time (sec)	Portfolio "Loss"	Portfolio Risk
0	0.00	0.10251679	10.46461509
10	8.86	0.10196217	10.06380700
20	9.12	0.10196221	10.06380700
30	8.98	0.10196221	10.06380700
40	8.85	0.10196221	10.06380700
50	9.03	0.10196221	10.06380700
60	9.02	0.10196221	10.06380700
70	8.92	0.10196221	10.06380700
80	9.00	0.10196221	10.06380700
90	9.06	0.10196221	10.06380700
100	8.75	0.10196221	10.06380700

Asset Class:	US Bonds	US Large-caps Equity	US Small-caps Equity	EAFE International Equity
Investment Allocation:	0.300000	0.30634937	0.130000	0.26365063

Portfolio Return: 9.8075561

Portfolio Risk: 10.06380700

Table 3C: GENO Input Data Matrix for the MV optimization problem with 'exposure' and 'round-lot' constraints

VARIABLE NAME	X1	X2	X3	X4
UCB	0.3	1	1	1
LCB	1e-6	1e-6	1e-6	1e-6
USB	1	1	1	1
LSB	1e-6	1e-6	1e-6	1e-6
Initial State Vector	0	0	0	0
Final State Vector	0	0	0	0
Discrete Values	0	0	0.13	0

Table 3D: Legend to the input data for the MV optimization problem with exposure and round-lot constraints

- Control variables (**u**): UCB — *U*pper *C*ontrol *B*ound; LCB — *L*ower *C*ontrol *B*ound;

- State variables (**x**): USB — *U*pper *S*tate *B*ound; LSB — *L*ower *S*tate *B*ound

- The entries 'UCB = 0.3' and 'Discrete Values = 0.13' implement the following constraints respectively:

  a) Exposure: Funds allocated to US Bonds (variable x_1) should not exceed 30% of the total fund;

  b) Round-lot: Investment in US Small-Caps Equity (variable x_3) should be in integer multiples of 0.13, say. [13]

Remarks on Example 3. This example shows how one may easily deal with important practical matters (such as the German 5-10-40 rule mentioned in §2) using GENO; the framework actually accommodates even more complex real-world scenarios that one may formulate, e.g. via disjunctive programming (see §5).

---

[13] Such a discrete investment proportion would have been derived as follows: let p and n denote the current market price and the round-lot quantity respectively for the asset in question. Then $c = np$ is the cost in currency units of one round-lot, and based on a total investment budget b, the discrete round-lot investment quantity on a per-unit-of-investment basis is simply c/b.

Example 4: A single-period MV optimization problem: The robust approach

MP11: $\quad \underset{\mathbf{u}}{\mathbf{Opt}} \left\{ m_g(\mathbf{u}),\ w_g(\mathbf{u}),\ m_h(\mathbf{u}),\ w_h(\mathbf{u}) \right\}$

Subject to: $\quad x_{k+1}^{j} = x_k^{j} + u_k^{j}$

$$\sum_{j=1}^{n} u_k^{j} = 1$$

$$u_k^{j} \in [0, 1]$$

$$\mathbf{x}_0 = \mathbf{0}$$

$$j \in \{1, 2, \ldots, 5\}$$

$$k \in \{0\}$$

Where: $\quad m_g(\mathbf{u}) \equiv 0.5(g_L(\mathbf{u}) + g_R(\mathbf{u}))$

$\qquad\qquad w_g(\mathbf{u}) \equiv 0.5(g_R(\mathbf{u}) - g_L(\mathbf{u}))$

$\qquad\qquad m_h(\mathbf{u}) \equiv 0.5(h_L(\mathbf{u}) + h_R(\mathbf{u}))$

$\qquad\qquad w_h(\mathbf{u}) \equiv 0.5(h_R(\mathbf{u}) - h_L(\mathbf{u}))$

Tables 4A: Expected Returns (%) [70]

Asset Name[14]	R1000 Growth	R1000 Value	R2000 Growth	R2000 Value	LBGC Bond
Mean Return (%)	15.1	12.5	14.7	9.02	17.68

Tables 4B: Covariance Matrix [70]

Covariance Matrix	R1000 Growth	R1000 Value	R2000 Growth	R2000 Value	LBGC Bond
R1000 Growth	2.8891	1.8417	3.2870	1.9204	0.1346
R1000 Value	1.8417	1.7427	2.3161	1.6879	0.1441
R2000 Growth	3.2870	2.3161	5.1551	3.0847	0.0859
R2000 Value	1.9204	1.6879	3.0847	2.4182	0.1158
LBGC Bond	0.1346	0.1441	0.0859	0.1158	0.1848

Notes on Example 4. This example affords an assessment of the robust optimization model presented in §**5** using data prepared by Tütüncü & Koenig [70, p.13]. But instead of solving it directly as a multi-objective problem, an *ad hoc* adaptation of the composite-metric method for solving nonlinear equation systems that is described in [63] was used for the sake of efficiency. A non-robust model was also run assuming nominal values for the expected return and the covariance matrix for comparison—both results are depicted below.

---

[14] Legend: **R1000 Growth** – Russell 1000 index for large cap growth stocks; **R1000 Value** – Russell 1000 index for large cap value stocks; **R2000 Growth** – Russell 2000 index for small cap growth stocks; **R2000 Value** – Russell 1000 index for small cap value stocks; **LBGC Bond** – Lehman Brothers US Intermediate Government/Credit Bond index for fixed income securities

GENO **Output** [Robust Model—Solution by Method EDR]

Generation Number	Time (sec)	d_c
0	0.00	5.52383467
50	14.87	5.02237228
100	14.79	5.02237228
150	14.85	5.02237228
200	14.84	5.02237228
250	14.80	5.02237227
300	14.77	5.02237226
350	14.85	5.02237223
400	14.90	5.02237223
450	14.90	5.02237223
500	14.90	5.02237223

Asset Class:	R1000 Growth	R1000 Value	R2000 Growth	R2000 Value	LBGC Bond
Allocation:	0.03958380	0.21741805	0.00000000	0.14103438	0.60196377

Portfolio Return:	0.894049342
Portfolio Risk:	0.649939001
Return-Risk Ratio:	1.375589609

GENO **Output** [Non-Robust Model]

Generation Number	Time (sec)	Portfolio "Loss"	Portfolio Risk
0	0.00	1.05423779	0.95834628
50	56.33	1.02718220	1.00101421
100	56.44	1.02717774	1.00100874
150	55.97	1.02717727	1.00100833
200	56.30	1.02717690	1.00100788
250	56.74	1.02717690	1.00100788
300	55.65	1.02717689	1.00100786
350	56.67	1.02717689	1.00100786
400	56.94	1.02717689	1.00100786
450	56.80	1.02716565	1.00099425
500	55.99	1.02716250	1.00099070

Asset Class:	R1000 Growth	R1000 Value	R2000 Growth	R2000 Value	LBGC Bond
Allocation:	0.170000000	0.461193750	0.017943750	0.050862500	0.300000000

Portfolio Return:	0.973555790
Portfolio Risk:	1.000990695
Return-Risk Ratio:	0.972592247

Remarks on Example 4. As can be seen above, the robust model generates a portfolio with much lower risk, but at the expense of lower returns; surprisingly, the robust portfolio is also less diversified than the non-robust version albeit marginally. But then again, this apparent "weakness" is more than compensated for by the significantly superior return-risk ratio, and in any case, the GENO framework allows one to easily induce more diversity via exposure constraints as illustrated by Example 3 above.

8 Discussion

Preamble. There currently exists various means for evaluating portfolios,[15] but most of the methods stem from an idea first suggested by Roy as a reasonable measure of an investment strategy's value, namely the 'Safety-First ratio' [57]. William Sharpe [60] applied Roy's ideas to the mean-variance framework of Markowitz and developed one of the best known performance evaluation measures in portfolio management now commonly known as the 'Sharpe ratio'. But this paper takes a different approach that is rooted in the notion that the asset allocation problem is fundamentally a multi-objective problem. The performance measure adopted is called the 'compromise metric' and it has already been explained in §6—it is simply the distance, in outcome space, from the solution point to the 'ideal outcome', the latter being that "win-win" situation in which each facet of the multi-objective allocation problem is optimized to its fullest extent within the operational constraints. In reading the tables of results to follow therefore, the compromise metric should be viewed as the primary criterion and all others as secondary; a portfolio with a lower value for the compromise metric should be preferred over that with a higher value; and quiet often (but not necessarily always) this ranking is replicated in the return-risk ratios (which are related to Sharpe ratios; see footnote **12**). Although the compromise metric seemingly lacks economic justification in the vein of the Sharpe ratio or its derivatives, it has substantive advantages over the latter: it is independent of the probability distribution of the asset returns; and its implicit assumption of rationality on the part of the model-user is consistent with choice theory.[16] Finally note that if a solution point is optimal in the compromise metric sense, then such a point necessarily lies on the Pareto frontier [73], and the said frontier is a realization of a well established multi-objective optimality concept called 'Pareto-dominance' that in fact subsumes the 'E-V rule' of Markowitz.

On Example 1. Results for this example are summarised in Table 5; portfolios generated by GENO are prefixed by the letter 'G'—the first set is by the 'sequential-decisions' model MP4; the second by the 'parallel-decisions' model MP5.

Table 5: A Comparative Analysis for **Example 1**

	G-PORTFOLIO 1	G-PORTFOLIO 2	L-PORTFOLIO 1	L-PORTFOLIO 2
Security 1	0.088790530	0.174000000	0.088000000	0.158000000
Security 2	0.250125980	0.171999970	0.251000000	0.155000000
Security 3	0.282629390	0.273000000	0.282000000	0.314000000
Security 4	0.103244630	0.081000030	0.104000000	0.038000000
Security 5	0.275209470	0.300000000	0.275000000	0.334000000
PORTFOLIO RETURN (%)	14.41893378	14.82511990	14.41178000	15.18698000
PORTFOLIO RISK	0.79053228	0.79942382	0.79053140	0.81108510
REWARD-RISK RATIO	18.23952574	18.54475636	18.23049660	18.72427443
COMPROMISE METRIC	0.79356862	0.80226452	0.79357076	0.81375348

---

[15] See e.g. [5, 21] for comprehensive accounts on existing methods of evaluating portfolios.

[16] Recall that the logic underlying the compromise solution concept is that: (i) if it were achievable, the ideal outcome would constitute *the* optimal solution to a multi-objective optimization problem; (ii) but since this is usually not the case, it is reasonable to posit that a rational decision-maker would accept a compromise "downwards" from the ideal to a less-than-ideal outcome as being the best that can be done under the circumstances provided the degree of the said "downward compromise" is minimal.

Remarks. These portfolios yield comparable performances. But according to the evaluation criterion adopted in this study, it is clear that G-PORTFOLIO 1 (which has the lowest compromise metric value at 0.79356862) is the best option. And in this instance, this rank is not reflected by the reward-risk ratios and so one would be committing an error if one were to rely on the Sharpe ratio as a measure of portfolio performance. Note also that portfolio G-PORTFOLIO 1 is only marginally poorer in terms of variance, than portfolio L-PORTFOLIO 1 that was computed by Luenberger [40] by explicitly minimizing the risk function.

On Example 2. The results reported by Michaud, Esch & Michaud [44] for this example that were presented earlier in Table 2C; the asset allocations are repeated below for convenience but the accompanying portfolio return and risk have been re-calculated using Microsoft Excel™; portfolios by GENO are prefixed by 'G'—the first set is by the 'sequential-decisions' model MP4; the second by the 'parallel-decisions' model MP5.

Table 6: A Comparative Analysis for **Example 2**

	G-PORTFOLIO 1	G-PORTFOLIO 2	MARKET PORTFOLIO	BLACK-LITTERMAN	MICHAUD MSR
Euro Bonds	0.24238574	0.5475000	0.2000000	0.2000000	0.2300000
US Bonds	0.00000000	0.0790000	0.2000000	0.2000000	0.1990000
Canada	0.00000000	0.0190000	0.0600000	0.0600000	0.0990000
France	0.07460059	0.3210000	0.0600000	0.0000000	0.0430000
Germany	0.00000000	0.0272500	0.0600000	0.0150000	0.0470000
Japan	0.19568359	0.0053000	0.0600000	0.0600000	0.0660000
UK	0.04040430	0.0009500	0.0600000	0.0150000	0.0540000
US	0.44692578	0.0000000	0.3000000	0.4500000	0.2620000
PORTFOLIO RETURN (%)	7.79632824	5.6859750	6.2800000	6.2095000	5.9737000
PORTFOLIO RISK (%)	11.27657900	9.9612040	9.6093113	9.6006700	9.3142055
REWARD-RISK RATIO	0.69137353	0.5708121	0.6535329	0.6467778	0.6413534
COMPROMISE METRIC	0.17078690	0.2021220	0.1859836	0.1874895	0.1915682

Remarks. The "Market Portfolio" which is required for the efficient re-sampling technique and the Black-Litterman model is arbitrarily defined by Michaud, Esch & Michaud [44] as comprising a

"60/40 asset mix of domestic and international stocks an d bonds with equal weights for the bond indices, equal weights for US versus non-US equity indices and equal weights for non-US indices" [paraphrased from p.9]

Although the revised figures for the other portfolios are better, a portfolio by GENO, namely G-PORTFOLIO 1, is once again the best option according to the compromise metric, and this time, the reward-risk ratio confirms this rank. Michaud, et al. consider their portfolio as superior because it is more diversified than the Black-Litterman portfolio, and that applies to G-PORTFOLIO 1 as well. But the question of what constitutes an 'optimally diversified' portfolio is far from settled [11, 64]; besides, this apparent "weakness" in the GENO method is more than compensated for by the superior return-risk ratio; and in any case, the GENO framework allows one to easily induce diversity via exposure constraints as shown by Example 3 above. Note also that G-PORTFOLIO 1 is independent of the "Market Portfolio" and G-PORTFOLIO 2 is almost as diversified as the Michaud MRS but sub-optimal, at least as far as the compromise metric and the reward-risk ratio are concerned.

On Example 3. The primary purpose of this example was to illustrate how one would implement 'exposure' and 'round-lot' constraints in the GENO framework. This has been sufficiently dealt with in §**5** above; the analysis and comments in this section pertain to the optimality of the GENO solution itself, i.e. we seek to establish whether GENO actually computes efficient portfolios.

It has already been pointed out that if a solution point is optimal in the compromise metric sense, then such a solution point necessarily lies on the Pareto frontier [73], and that Pareto-efficiency subsumes the 'E-V rule'. And so in order to prove whether a given portfolio is efficient, one only needs to show that the return-risk outcome said portfolio entails is indeed closest to the ideal. To that end, one approach is to generate other portfolios "near-by" and simply measure and compare the corresponding distances to the ideal outcome.

The table below presents a simple "perturbation analysis" that was performed using a spreadsheet program. The original solution is labelled G-PORTFOLIO, and PORTFOLIOS 1 – 4 are minor variants of it which were derived by slightly altering the allocation for each asset in turn, and then adjusting the remaining allocations accordingly so as to ensure that the new portfolio is feasible, and that its outcome remains on the same Pareto frontier.[17]

Table 7: A Perturbation Analysis of the GENO solution to **Example 3**

	G-PORTFOLIO	PORTFOLIO 1	PORTFOLIO 2	PORTFOLIO 3	PORTFOLIO 4
US Bonds	0.3000000	0.2500000	0.3000000	0.3000000	0.3000000
US Large-caps Equity	0.3063520	0.3063138	0.0063100	0.1763100	0.1063100
US Small-caps Equity	0.1300000	0.1300000	0.1300000	0.2600000	0.1300000
EAFE International Equity	0.2636480	0.3136862	0.5636900	0.2636900	0.4636900
PORTFOLIO RETURN (%)	9.8081774	10.0625830	10.0175830	9.9505830	9.9475830
PORTFOLIO RISK (%)	10.0638070	10.6562468	11.2958708	10.4515806	10.6291140
REWARD-RISK RATIO	0.9745991	0.9442893	0.8868358	0.9520649	0.9358807
COMPROMISE METRIC	10.0638070	10.6562468	11.2963119	10.4506380	10.6291140

Remarks. As can be seen from the row headed 'Compromise Metric', the G-PORTFOLIO is indeed closest to the ideal point in outcome space and it is therefore Pareto-efficient. Note that although the analysis was limited to a few points in a small neighbourhood of the original solution, it would be safe to assume that the G-PORTFOLIO is the unique minimum-distance-point to the ideal outcome. This follows because Pareto frontiers are normally "convex to the left" [40, p.156], and as such, they are tangential at only one point to a zero-centred circle in outcome space. And in this case the G-PORTFOLIO is also optimal in another sense—it has the highest reward-risk ratio as compared to all the "near-by" portfolios with outcome points on the same Pareto frontier.

---

[17] The test for checking whether the newly generated portfolio were Pareto-efficient is based on the parametric representation of Pareto frontiers, viz.:

$$\text{PF} = \lambda g(\mathbf{u}) + (1 - \lambda)h(\mathbf{u})$$

The requirement, in terms of variations, is that $\delta\text{PF} = 0$; thus, the test comprised checking the truth value of the following statement

$$\exists \lambda \in (0, 1) : \lambda\delta g(\mathbf{u}) + (1 - \lambda)\delta h(\mathbf{u}) \equiv 0$$

An alternative method for generating efficient portfolios that are "near-by" would be to use the two-fund theorem which asserts that

"Two efficient funds (portfolios) can be established so that any efficient portfolio can be duplicated, in terms of mean and variance, as a combination of these two; all investors seeking efficient portfolios need only invest in these two funds" [40, p.163]

This is the method by which Luenberger [40, p.164] generated the second portfolio labelled L-PORTFOLIO 2 in Example 1 above.

On Example 4. This example was primarily meant to evaluate the robust model presented in §**5**; a secondary purpose was to assess the relative merits of two other methods of computing portfolios—the said methods are adaptations of the 'NCP' (code: ECR) and 'Composite-metric' (code: EDR) methods for solving nonlinear equation systems that are explained in detail in [63]. Table 8A and 8B are similar to those presented earlier and their contents should be self-explanatory; Table 8C on the other hand is unique—it depicts variations in portfolio evaluators shown in the header row as a result of perturbations in the input data; it is an "opportunity cost" analysis performed via a spread-sheet program assuming the optimal portfolios generated by each method; each element of the expected returns vector was in turn assigned its 2.5 and 97.5 percentile value and the portfolio evaluators re-calculated; the figures reported are the net difference in evaluator values from their nominal values expressed as percentages of the later—a positive number indicates an increase. Note that this is not an evaluation of the robust model vis-à-vis the stability issues mentioned in §**2** because, provided the intervals in **d** are wide enough to accommodate future realizations of returns, it would be reasonable to posit stable portfolios. And so the aim here is rather different: it is to simulate and examine the "opportunity costs" of holding a supposedly optimal portfolio, based as it is on a particular forecast of asset returns and their correlation, when the said forecast does not actually pan out as expected, which is usually the case in practice.

Table 8A: A Comparison of solutions [Non-Robust vs. Robust Portfolios]

		PORTFOLIO RETURN	PORTFOLIO RISK	COMPROMISE METRIC	RETURN-RISK RATIO
NON-ROBUST MV MODEL	PORTFOLIO 1 [MO]	0.973555790	1.000990695	1.434240278	0.972592247
	PORTFOLIO 2 [ECR]	0.819166675	0.476734565	1.310539282	1.718286724
	PORTFOLIO 3 [EDR]	0.819167188	0.476735330	1.310538848	1.718285044
ROBUST MV MODEL	PORTFOLIO 4 [MO]	0.885258010	0.692700272	1.325089378	1.277981322
	PORTFOLIO 5 [ECR]	0.894048993	0.649938070	1.293629491	1.375591052
	PORTFOLIO 6 [EDR]	0.894049342	0.649939005	1.293629584	1.375589609

Table 8B: Composition of Portfolios [Non-Robust vs. Robust Portfolios]

	NON- ROBUST MV MODEL			ROBUST MV MODEL		
	PORTFOLIO 1	PORTFOLIO 2	PORTFOLIO 3	PORTFOLIO 4	PORTFOLIO 5	PORTFOLIO 6
R1000 Growth	0.170000000	0.000000000	0.000000000	0.103000000	0.039583690	0.039583800
R1000 Value	0.461193750	0.090312660	0.090313460	0.041800000	0.217417600	0.217418050
R2000 Growth	0.017943750	0.000000000	0.000000000	0.015109900	0.000000000	0.000000000
R2000 Value	0.050862500	0.095808810	0.095809320	0.240090100	0.141034010	0.141034380
LBGC Bond	0.300000000	0.813878530	0.813877220	0.600000000	0.601964700	0.601963770
PORTFOLIO RETURN	0.973555790	0.819166675	0.819167188	0.885258010	0.894048993	0.894049342
PORTFOLIO RISK	1.000990695	0.476734565	0.476735330	0.692700272	0.649938070	0.649939005
REWARD-RISK RATIO	0.972592247	1.718286724	1.718285044	1.277981322	1.375591052	1.375589609
COMPROMISE METRIC	1.434240278	1.310539282	1.310538848	1.325089378	1.293629491	1.293629584

Table 8C: Net Percentage Changes in Portfolio Evaluators

		PORTFOLIO RETURN (%)	COMPROMISE METRIC (%)	RETURN-RISK RATIO (%)
NON-ROBUST MV MODEL	PORTFOLIO 1 [MO]	-0.1337002	0.9516442	-0.1337002
	PORTFOLIO 2 [ECR]	-0.1087518	1.0951092	-0.1087518
	PORTFOLIO 3 [EDR]	-0.1087524	1.0951084	-0.1087524
ROBUST MV MODEL	PORTFOLIO 4 [MO]	-0.1643917	0.8114607	-0.1643917
	PORTFOLIO 5 [ECR]	-0.1386964	0.8160602	-0.1386964
	PORTFOLIO 6 [EDR]	-0.1386967	0.8160608	-0.1386967

Remarks. Admittedly, the analysis is limited in that only a few perturbations are considered, and all within the assets returns data (although, in mitigation, one may cite Chopra & Ziemba [9] who show that it is errors in estimated returns that have the most pronounced effects); secondly, the perturbations are applied in isolation and never as combinations, as would most likely be the case in reality. Accordingly, the following comments should not be treated as definitive.

Table 8B indicates that, within each class of model, the ECR and EDR portfolios are less diversified than the MO portfolios that are computed as per the original multi-objective specifications in MP4 and MP5. And depending on the type of norm is used to compute the compromise metric, the diversity of a portfolio may be a significant factor of the said metric. And so in order to avoid an 'apples-and-oranges' type of error, comparative remarks across the Robust / Non-Robust class divide shall concern portfolios computed by the same methods, i.e. PORTFOLIO 1 vs. PORTFOLIO 4, PORTFOLIO 2 vs. PORTFOLIO 5 and PORTFOLIO 3 vs. PORTFOLIO 6.

Judging by the compromise metric, Table 8A suggests that, within either class, portfolios computed by the ECR and EDR methods are better than those by the MO methods; and across the class divide, portfolios generated by robust models are better than their non-robust counterparts. Note that if the return-risk ratio (and hence the Sharpe ratio) were the primary 'goodness' criterion, the intra-class ranks would be confirmed in both cases, but not across the class divide.

As regards Table 8C, a natural tendency would be to focus exclusively on the potential loss signalled by the percentage changes in portfolio returns. And ranking the portfolios on that basis, it would appear that the ECR and EDR portfolios are likely to suffer lower losses than their MO counterparts as a result of perturbations in the input data. But such a ranking would sub-optimal because the risk associated with each potential loss (as represented by the portfolio variance) is not taken into account. A better approach would be to base one's decision on the compromise metric since this subsumes changes in both the portfolio return and portfolio risk; and in that case, Table 8C shows that, within each class, portfolios by ECR and EDR methods are likely to "cost" more than those by MO methods; and across classes, portfolios by robust models are likely to "cost" less than their non-robust counterparts following unexpected outcomes in the returns data.

In summary, one may conclude as follows: although based on a limited set of perturbation trials, the results indicate the relative superiority of the robust approach over the non-robust model; but within each class of models, apart from the fact that the two uni-objective methods (ECR and EDR) obviously take less time, there seems to be no consistent advantage in using ECR or EDR methods of computation over the original multi-objective compromise approach; the uni-objective methods tend to generate "less diverse" portfolios.

9 Summary and Conclusions

This paper has reviewed the portfolio model of Harry M. Markowitz [42], and its main aim was to present conceptual and concrete methods (in the context of an evolutionary solver called GENO) that effectively address practical issues that apparently impede its implementation in practice—a phenomenon known as the 'Markowitz optimization enigma' [46]. To that end, two generic models that are easily implemented were presented, namely (i) a pseudo-dynamic sequential-decisions model for single-period asset allocation; (ii) a truly dynamic parallel-decisions model that is primarily meant for multi-period asset allocation but equally applies to the single-period setting. Both models retain the multi-attribute nature of the portfolio problem: the basic criterion set comprises two measures of performance namely, the portfolio risk (for which the GENO framework allows any risk measure), and the portfolio reward or a proxy thereof; for either model, the criterion set may be augmented by extra objectives designed to address specific aspects of the 'Markowitz optimization enigma'.

It has been shown how *ad hoc* constraints such as those regarding 'exposure', 'cardinality', 'round-lots' and 'buy-in thresholds', may easily be incorporated into the basic model, sometimes by merely entering the required numbers into the input data matrix supplied to GENO; but for more complex scenarios, the required constraints may be implemented using binary variables, or via Boolean variables in conjunction with logic propositions—the final problem specification in either case is a mixed-variable program for which GENO has the requisite capacity; see numerical examples in [61].

It has also been shown how uncertainty in the model parameters may be accommodated using techniques from the calculus of real intervals; a tractable multi-objective model of the portfolio problem that explicitly accounts for data uncertainty was derived and its essence (with respect to the said uncertainty) was shown to be of the mean-risk variety; the robust formulation may easily be applied to portfolio models that involve more sophisticated risk measures (such as the ICTE in equation 1) provided the uncertain model parameters in the said measures feature in functions that are monotonic with respect to the said parameters [48, §5.2].

The compromise solution was deemed most appropriate for the models presented in §3. The compromise concept defines a single-point on the Pareto-efficient frontier, and the associated algorithm is more utilitarian because it dispenses with the need to calibrate utility functions as is typical of other single-point MV solution methods—see e.g. [15]; the model-user is not even required to express any preference prior, post or during the solution process. Furthermore, unlike many multi-objective methods that "tend to become cumbersome and even useless as we increase the number of objectives" [10, p.4], the size of the problem (as measured by the dimension of the outcome space) is irrelevant because one can always resort to the more efficient ECR and EDR computational methods if need be, and so the method is well suited to extended portfolio models such as those advocated in [28, 67, 68].

Several numerical examples have been presented and evaluated on the basis of the compromise metric; the results show that not only does GENO generate efficient solutions to the portfolio optimization problem in a single run, in the majority of cases, its solutions are also optimal in a much wider sense, namely that, as compared to all other points on the Pareto frontier, they have the highest return-risk ratio—a proxy indicator for 'value-for-money'; this is most apparent in Example 2 where, as compared to the patented efficient re-sampling method, the GENO portfolio is admittedly less diversified but has a far superior return-risk ratio. The robust version of the MV model is not only easy to implement, but it is also superior to its non-robust counterpart in terms of both nominal performance and the potential for avoiding "opportunity costs" due data uncertainty, at least according to the perturbation analysis performed on Example 4, albeit limited.

10 Legalities

I. Licence and Trademarks

Except for the trademark items mentioned below, this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit this page;[18] to override specific prohibitions of the governing licence,[19] submit a formal request to the author at: ike_siwale@hotmail.com

GENO™ is a trademark of Apex Research Ltd
Copyright © 1997-2015
All Rights Reserved Worldwide

GAUSS™ is a trademark of Aptech Systems Inc.
Copyright © 1983-2015
All Rights Reserved Worldwide

II. Disclaimer

This document contains proprietary material created by Apex Research Ltd which is subject to further verification and change without notice; however, Apex Research Ltd is under no obligation to provide an updated version. Furthermore, Apex Research Ltd does not make any expressed or implied warranty— including the warranties of merchantability and fitness for a particular purpose—as to the accuracy or completeness of the methods described herein; accordingly, Apex Research Ltd accepts no liability for any damages that may occur from use.

---

[18] http://creativecommons.org/licenses/by-nc-sa/4.0/

[19] http://creativecommons.org/licenses/by-nc-sa/4.0/legalcode

Appendix A

The Markowitz Mean-Variance Criterion

Consider a portfolio of assets whose return on investment, r, is random variable characterised by a probability distribution with mean μ and standard deviation σ, and a utility function $\mathbb{U}$ that depends on the portfolio return *r* and possibly other factors. A Taylor series expansion of $\mathbb{U}$ around the population mean μ is given by

$$\mathbb{U}(r) = \sum_{n=0}^{\infty} \frac{(r-\mu)^n}{n!} \cdot \frac{\partial^n \mathbb{U}(r)}{\partial r^n}\bigg|_{r=\mu} \qquad (A1)$$

Utility Theory asserts that investors are 'utility maximizers', but before one can maximise one's utility, one must know the parameters of the return distribution. Now it is well known that if a utility maximizer evaluates a probability distribution on the basis of the first *n* moments, then his or her von Neumann-Morgenstern utility function is an *n*-th degree polynomial; conversely, if investor's utility function is an *n*-th degree polynomial, then he or she evaluates probability distributions based only on the first *n* moments [49]. Therefore in treating the investor as a 'Mean-Variance optimizer', we need only consider a truncated Taylor series approximation of his utility function, namely the second-degree polynomial:

$$\mathbb{U}(r) = \mathbb{U}(\mu) + \mathbb{U}'(\mu)(r-\mu) + 0.5\mathbb{U}''(\mu)(r-\mu)^2 \qquad (A2)$$

Applying the expectation operator $\mathbb{E}$ to both sides of equation (A2) leads to the expected utility equation:

$$\mathbb{E}[\mathbb{U}(r)] = \mathbb{U}(\mu) + \mathbb{U}'(\mu)\mathbb{E}[(r-\mu)] + 0.5\mathbb{U}''(\mu)\mathbb{E}[(r-\mu)^2] \qquad (A3)$$

But by definition of the parameters μ and σ, we have that:

$$\mu = \mathbb{E}[r] \quad \Rightarrow \quad \mathbb{E}[(r-\mu)] = 0 \qquad (A4)$$

$$\sigma^2 = \mathbb{E}[(r-\mu)^2] \qquad (A5)$$

Hence:
$$\mathbb{E}[\mathbb{U}(r)] = \mathbb{U}(\mu) + 0.5\sigma^2\mathbb{U}''(\mu) \qquad (A6)$$

Consider an investor's iso-utility contours (a.k.a. 'indifference curves') in the Mean-Variance plane. Along any such contour, the expected utility is constant and therefore its total differential is identically zero, i.e.:

$$d\mathbb{E}[\mathbb{U}(r)] = \mathbb{U}'(\mu)d\mu + 0.5\{\mathbb{U}''(\mu)d(\sigma^2) + \sigma^2\mathbb{U}'''(\mu)d\mu\} \equiv 0 \qquad (A7)$$

But by assumption, $\mathbb{U}$ is a second degree polynomial and therefore $\mathbb{U}'''(\mu) \equiv 0$; and upon factoring out the term $U'(\mu)$ from the RHS of Equation (A7), we have that:

$$d\mathbb{E}[(\mathbb{U}(r))] = \mathbb{U}'(\mu)\left\{d\mu + 0.5\frac{\mathbb{U}''(\mu)}{\mathbb{U}'(\mu)}d(\sigma^2)\right\} \equiv 0 \qquad (A8)$$

Furthermore, since $\mathbb{U}'(\mu) \neq 0$ (because otherwise, the utility function would be independent of portfolio return, contrary to the basic tenet of Utility Theory itself), it follows immediately that:

$$d\mathbb{E}\big[\mathbb{U}(r)\big] \equiv d\mu + 0.5 \frac{\mathbb{U}''(\mu)}{\mathbb{U}'(\mu)} d(\sigma^2) = 0 \tag{A9}$$

Integrating both sides of Equation (A9) and introducing the factor $\lambda = -\mathbb{U}''(\mu)/\mathbb{U}'(\mu)$ —known as the 'Arrow-Pratt Measure of Absolute Risk Aversion', or simply the 'Coefficient of Absolute Risk Aversion'—we have that:

$$\mathbb{E}\big[\mathbb{U}(r)\big] = \mu - 0.5\lambda\sigma^2 \tag{A10}$$

But the population parameters, μ and σ, would normally be unknown to the investor: for practical purposes, one has to replace these by their respective estimates. Typically, given a sample of portfolio returns, the "plug-in" solution strategy advocates replacing μ by the sample mean, r, and σ^2 by the sample variance, s^2.

$$\mathbb{E}\big[\mathbb{U}(r)\big] = r - 0.5\lambda s^2 \tag{A11}$$

But often this Mean-Variance criterion is presented in terms of the individual assets compromising the portfolio. Thus, for an *n*-asset portfolio, we have that:

$$\mathbb{E}\big[\mathbb{U}(r)\big] = \sum_{i=1}^{n} u_i r_i - \varphi^{-1} \sum_{i=1}^{n} \sum_{j<i} u_i u_{ij} c_{ij} \tag{A13}$$

Where: r^i is the expected return for the *i*-th asset;

c^{ij} is the covariance of returns between the *i*-th and *j*-th assets;

x^i is the proportion of wealth invested in the *i*-th asset;

φ is the investor's risk tolerance (Note: $\varphi = 2\lambda^{-1}$).

Remarks. The Mean-Variance criterion is seemingly based on a sound theoretical base however problems emerge in the implementation phase. The main source of difficulty is in calibrating the risk aversion coefficient λ (or the risk tolerance φ), and unfortunately, economic theory isn't much help because there is no consensus amongst economists themselves as to what these parameters entail and how one should go about measuring them in practice. For example, Mathew Rabin [54] writes:

"Expected utility theory seems to be a useful and adequate model of risk aversion for many purposes, and it is especially attractive in lieu of an equally tractable alternative model [. . . but the theory] makes wrong predictions about the relationship between risk aversion over modest stakes and risk aversion over large stakes. Hence, when measuring risk attitudes according to the expected utility hypothesis, differences in estimates may come from differences in the scale of risk comprising the data sets rather than from differences in risk attitude of people being studied [. . .] so not only are standard measures of risk aversion somewhat hard to interpret [. . .] but even attempts to compare risk attitudes across groups will be misleading unless economists pay due attention to the theory's calibration problems" [Paraphrased from p.1282 & p.1287]

The critical-line algorithm and the parametric approach are of course practical 'work-arounds' that have been used since the inception of the MV decision model with some measure of success. The remarks above further underline the practicality of the multi-objective approach present here: the decision model does not involve an artificial scalarization of the criteria, and one does not need the coefficient of absolute risk aversion to solve it.

Appendix B

The Iterated CTE: A Dynamic Risk Measure

I. Introduction

A Risk Measure is a function that assigns a real number to the outcome of a random process. An overview and pedagogical introduction to the subject of risk measurement may be found in [13, 69].

Risk measures may be classified into two broad categories:

- Type I: Measures that compute the deviation of a random outcome from a target value (often the Mean);

- Type II: Measures that compute potential loss of a random outcome with a pre-specified confidence level.

Risk measures of the first type include (i) two-sided deviations from a relevant target value, e.g. Variance; (ii) one-sided 'shortfalls' relative a relevant target value, e.g. Lower Partial Moments. This Appendix is only concerned with measures of the second type; we begin with 'Value-at-Risk' which is defined as follows.

Consider a *positive* random variable, L; a probability measure P; a confidence level $\alpha \in (0, 1)$; and a time interval $\tau = [0, T]$. The Value-at-Risk over τ is defined as the smallest number λ such that the probability that the loss variable L at time T exceeds λ is no greater than $(1 - \alpha)$, i.e.:

$$\text{VaR}_\alpha(L) \equiv \text{Inf}\{\lambda \in \mathbb{R} : P(L > \lambda) \leq 1 - \alpha\} \tag{B1}$$

Typical values of the confidence parameter are $\alpha = 0.95$ and $\alpha = 0.99$; in market risk management the time horizon τ is usually one or ten days; in credit and operational risk management, τ is usually taken as one year. Equation (B1) says that in $100\alpha\%$ of its realisations, the loss variable L is less than or equal to $\text{VaR}_\alpha(L)$; conversely, the loss L exceeds $\text{VaR}_\alpha(L)$ only in $100(1 - \alpha)\%$ of the outcomes; in terms of capital requirement, the equation says that a reserve of $\text{VaR}_\alpha(L)$ would be sufficient to cover losses in 100α cases out of 100. Value-at-Risk is, in essence, a quantile measure, and one may formally demonstrate this as follows:

- DEFINITION B1 [Generalised Inverse]: Given a *non-decreasing* function $\phi : \mathbb{R} \to \mathbb{R}$, the generalized inverse of ϕ is defined by $\phi^{\leftarrow}(y) = \text{Inf}\{x \in \mathbb{R} : \phi(x) \geq y\}$, where we use the convention that the infimum of an empty set is ∞.

- DEFINITION B2 [Quantile Function]: Given a cumulative distribution function F, the generalized inverse denoted by $F^{\leftarrow}(\cdot)$ is called a quantile function of F; and for $\alpha \in (0, 1)$, the α-quantile of F is given by

$$q_\alpha(F) \equiv F^{\leftarrow}(\alpha) = \text{Inf}\{\lambda \in \mathbb{R} : F(\lambda) \geq \alpha\}$$

If F is *strictly increasing* then we simply have $q_\alpha(F) = F^{-1}(\alpha)$, where $F^{-1}(\cdot)$ denotes the ordinary inverse of F.

- LEMMA B1 [α-quantile]: A point $\lambda_0 \in \mathbb{R}$ is the α-quantile of the cumulative distribution function F if and only if the following two conditions hold for all $\lambda < \lambda_0$: (i) $F(\lambda_0) \geq \alpha$; (ii) $F(\lambda) < \alpha$.

The lemma follows immediately from the definition and the right-continuity of F. And by definition of a cumulative distribution function, Equation (B1) implies that:

$$\text{VaR}_\alpha(L) \equiv \text{Inf}\{\lambda \in \mathbb{R} : 1 - F_L(\lambda) < 1 - \alpha\} \tag{B2a}$$

Equation (B2a) together with DEFINITION B2 leads to:

$$VaR_\alpha(L) \equiv Inf\{\lambda \in \mathbb{R} : F_L(\lambda) \geq \alpha\} \equiv q_\alpha(F_L) \tag{B2b}$$

Although, as a risk measure, VaR satisfies several 'goodness' criteria, it has some serious drawbacks that many have pointed out: by its very definition, VaR is only concerned with the frequency of losses but not the size of the loss;[20] and much more importantly, it is not universally coherent [16, 17, 18]. It was this failure of VaR to comply with the compelling theory of coherency developed in [3, 4] that led to the development of a risk measure called 'Conditional Tail Expectation'.

II. Conditional Tail Expectation (CTE)

This risk measure is defined as follows.

▪ DEFINITION B3A [CTE]: Let F be a *strictly increasing* cumulative distribution function; let $VaR_\alpha(L)$ denote the Value-at-Risk for a loss random variable L with respect to F and at confidence level α; and let $\mathbb{E}$ denote the mathematical expectation operator. Then the conditional tail expectation (CTE) of L is defined as:

$$CTE_\alpha(L) \equiv \mathbb{E}\big[L \,|\, L > VaR_\alpha(L)\big] \tag{B3a}$$

▪ REMARK B1: Note that this definition implicitly assumes that $VaR_{\alpha+\varepsilon}(L) > VaR_\alpha(L)$ for all $\varepsilon > 0$. To cover the possibility of a fall in probability mass, i.e. when $\exists\varepsilon > 0 : VaR_{\alpha+\varepsilon}(L) = VaR_\alpha(L)$, a more general formulation is required, and this is as follows.

▪ DEFINITION B3B [CTE]: Let F be a cumulative distribution function; let $VaR_\alpha(L)$ denote the Value-at-Risk for a loss random variable L with respect to F and at confidence level α; let $\mathbb{E}$ denote the mathematical expectation operator; and let the parameter β be defined as follows:

$$\beta \equiv Max\{\lambda : VaR_\alpha(L) = VaR_\lambda(L)\}$$

Then the conditional tail expectation of L is defined as:

$$CTE_\alpha(L) \equiv \frac{(1-\beta)\mathbb{E}\big[L \,|\, L > VaR_\alpha(L)\big] + (\beta-\alpha)VaR_\alpha(L)}{1-\alpha} \tag{B3b}$$

An alternative expression for CTE_α that dispenses with the auxiliary parameter β is as follows:

$$CTE_\alpha(L) \equiv VaR_\alpha(L) + \frac{P\big(L > VaR_\alpha(L)\big)}{1-\alpha}\,\mathbb{E}\big[L - VaR_\alpha(L) \,|\, L > VaR_\alpha(L)\big] \tag{B3c}$$

▪ REMARK B2: The last definition clearly shows that CTE will always lead to a risk level that is at least as high as Value-at-Risk. Loosely speaking, the CTE may be described as the 'mean loss in the worst $100(1 - \alpha)\%$ realisations of the loss random variable L'.

Since the CTE is a single-period risk measure, it may not be suitable in some cases as Wang explains [75]:

1. In many applications, the risks are multi-period in nature due to intermediate cash flows: an example would be a position in futures which results in intermediate cash flows due to the accounting practice of 'marking-to-market';

2. Whereas regulators require financial institutions to demonstrate adequate capital 'for the next ten days', most financial institutions normally calculate their VaRs on a daily basis, and each calculation pertains to a different 10-day period; this 'moving horizon problem' requires risk measures that are, in some sense, consistent over time;

3. Risk management is often part of a larger optimization problem such as portfolio management, and if the optimization problem is dynamic, then to ensure time-consistency, the risk measure must also be dynamic.

---

[20] Rockafellar explains: "Inadequate account is taken of the degree of danger inherent in potential violations beyond [the VaR] level. In cases where [L > VaR] which do occur with probability (1 - α), is there merely an inconvenience or a disaster?" [Paraphrased from 56, p.42]

In dynamic optimisation, 'time-consistency' is an important attribute of the solution. Consider a situation where a solution has been computed and actually implemented on $[0, t) \subset [0, T]$. If a re-calculation at time t yields the same solution for the remainder of the optimisation period $[t, T]$ as originally determined prior to implementation, then such a solution is said to be time-consistent [37]. This property is in essence a statement of the Richard Bellman's 'Principle of Optimality' which states:

> "An optimal policy has the property that whatever the initial state and decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [7, p.83].

Unfortunately, the literature on dynamic risk measures is currently rather limited, and the few models that are available, such as the formulations in Wang [75], Riedel [55] and Artzner, et al. [2], are difficult to implement in practice. According to Hardy and Wirch:

> "The key to defining [the risk measures of Wang, Riedel and Artzner, et al.] is determining the closed convex set of test probabilities (scenarios) over which to calculate the expected values; defining this set is the most difficult hurdle to implementing multi-period risk measures" [Paraphrased from 26 p.63].

Their alternative approach is based on the premise that, in the multi-period setting, the appropriate value to consider in assessing capital adequacy is not the final liability, but the periodic amounts that are required 'per decision stage'—each being the amount determined by the risk measure as 'necessary to cover possible losses in the next time period'. They call their measure the 'Iterated Conditional Tail Expectation'.

III. The Iterated Conditional Tail Expectation (ICTE)

The iterated conditional tail expectation of Hardy & Wirch [26] employs a result that states that any risk measure can be expressed as the expected value of a loss random variable under a suitably distorted probability measure [74, 30]; this result shall be assumed in this exposition.

Figure B1: Multi-period Risk assessment of a Liability *L* through *n* Iteration Intervals each of length τ years
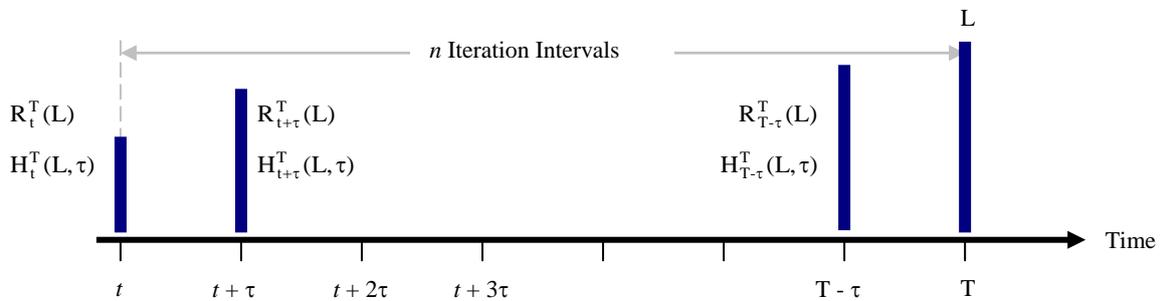
The timeline in Figure B1 is a useful visual aid; in the diagram, 'now' is labelled as time t, and the planning horizon $[t, T]$ is divided into n time intervals, each of length τ years. This assumption of equal length time steps is not necessary—it is merely for analytical expediency. Let H and R denote two risk measures and let L be a liability due at time T. Let the capital required to meet the said liability (as determined at time t by R) be $R_t^T(L)$; at the same time, let $H_t^T(L, \tau)$ denote the same capital requirement but this time as determined by H which is computed by backward recursion from T down to t in steps of length τ. Let $\mathbb{E}_D[\cdot]$ denote the expectation operator under some distorted probability distribution function, and let $\mathbb{F}_t$ denote the sigma algebra representing information that is pertinent at any time t; assume a nominal rate of δ for continuous discounting.

At time t, the required capital may be stated thus:

$$R_t^T(L) \equiv \mathbb{E}_D\left[Le^{-\delta n\tau} \mid \mathbb{F}_t\right] \tag{B4a}$$

And in terms of the risk measure H—which is a functional recursion by design—we have that:

$$H_t^T(L, \tau) \equiv \mathbb{E}_D\left[H_{t+\tau}^T(L, \tau)e^{-\delta t} \mid \mathbb{F}_t\right] \tag{B4b}$$

The right hand side of Equation (B4b) may be expressed in terms of the measure R as follows:

$$\mathbb{E}_D\left[H_{t+\tau}^T(L, \tau)e^{-\delta k} \mid \mathbb{F}_t\right] \equiv R_t^{t+\tau}\left(H_t^T(L, \tau)\right) \tag{B4c}$$

Hence:

$$H_t^T(L, \tau) = R_t^{t+\tau}\left(H_{t+\tau}^T(L, \tau)\right) \tag{B4d}$$

A repeated application of the logic underlying B4b and B4c on $H_{t+\tau}^T$, $H_{t+2\tau}^T$, $H_{t+3\tau}^T$, etc readily shows that Equation (B4d) written in expanded form is as follows:

$$H_t^T(L, \tau) = R_t^{t+\tau}\left(R_{t+\tau}^{t+2\tau}\left(R_{t+2\tau}^{t+3\tau}\left(\cdots R_{T-\tau}^T\left(H_T^T(L, \tau)\right)\right)\right)\right) \tag{B4e}$$

But by definition, $H_T^T(L, \tau) = L$; and so we have that:

$$H_t^T(L, \tau) = R_t^{t+\tau}\left(R_{t+\tau}^{t+2\tau}\left(R_{t+2\tau}^{t+3\tau}\left(\cdots R_{T-\tau}^T(L)\right)\right)\right) \tag{B5}$$

Equation (B5) says that at time t, the capital determined by H is a result of a backward recursive application of the measure R using the capital requirement at the next 'forward' iteration stage as the argument. Note that the functional form of R is still undeclared: if one defines this as the 'conditional tail expectation of its respective argument', then by Definition B3a and using the notation $\theta_t \equiv H_t^T(L, \tau)$, the right hand side of B4d is:

$$R_t^{t+\tau}\left(\theta_{t+\tau}\right) \equiv \mathbb{E}_D\left[\theta_{t+\tau} \mid \theta_{t+\tau} > q_\alpha\left(\theta_{t+\tau}\right)\right] \tag{B6a}$$

Now, at time $t = T - \tau$, we have that:

$$\left.H_t^T(L, \tau)\right|_{t=T-\tau} = \left.R_t^{t+\tau}\left(\theta_{t+\tau}\right)\right|_{t=T-\tau} \tag{B6b}$$

Replacing the index t by $T - \tau$ yields:

$$H_{T-\tau}^T(L, \tau) = R_{T-\tau}^T\left(\theta_T\right) \tag{B6c}$$

But by definition, $\theta_T \equiv H_T^T(L, \tau) = L$, hence we have that:

$$H_{T-\tau}^T(L, \tau) = R_{T-\tau}^T(L) \equiv \mathbb{E}_D\left[L \mid L > q_\alpha(L)\right] \tag{B6d}$$

Thus a definition of the one-step risk measure R induces a similar functional form for the recursive risk measure H. This observation prompted Hardy and Wirch to assert:

> "Any single period risk measure may be used to construct a multi-period risk measure by iteration. If we start with a risk measure having some attractive features, such as the CTE, then we retain most of those features after iteration" [Paraphrased from 26, p.65]

Wirch and Hardy provide proof of the above assertion: they demonstrate that, just like the 'ordinary' conditional tail expectation, the *iterated* conditional tail expectation satisfies the static coherency axioms of Artzner, et al. [3, 4] at each stage of iteration. But in addition, the ICTE also satisfies two dynamic axioms proposed by Riedel [55] called 'dynamic consistency' and 'relevance'; and in most practical cases, the ICTE also has a property they call 'Time-step Increasing'. These concepts are defined as follows.

- DEFINITION B4A [Dynamic Consistency]:   A dynamic risk measure is said to be 'dynamically consistent' if, for any two random loss process X and Y, and assuming no cash flows in the interval [t-s, t], we have that:

$$H_t^T(X,k) \mid \mathbf{F}_t \equiv H_t^T(Y,k) \mid \mathbf{F}_t, \forall \mathbf{F}_t \quad \Rightarrow \quad H_{t-s}^T(X,k) \equiv H_{t-s}^T(Y,k), \forall s \in (0,t]$$

- REMARK B3: In other words, 'dynamic consistency' requires that if risk measures for the losses X and Y are the same for every possible probabilistic outcome at time *t*, they must be the same for all previous dates provided there are no cash flows in the interim period. 'Dynamic Consistency' is also called 'Time Consistency'.

- DEFINITION B4B [Relevance]:   A dynamic risk measure is said to be 'relevant' if every loss which is not excluded by the current history carries positive risk.

- DEFINITION B4C [Time-step Increasing]: A risk measure *H* is said to be 'time-step increasing' if, for a liability *L* due at time *T*, the risk assessed at earlier time *t* is an increasing function of the number of iteration stages *nk* into which the interval [t, T] is sub-divided (see Figure B1).

- REMARK B4:   This property is intuitively appealing—it essentially states that 'the more frequent one checks for potential problems, the larger the probability of detecting them'. If the time-step property holds for the risk measures *H* and *R* described above, then it means that $H \geq R$ since *R* is a single time-step version of *H*.

Hardy & Wirch [26] derive a closed-form expression for the ICTE in the case where the loss random variable is a log-normal process. Their derivation is presented in the following section.

IV. ICTE: The Log-Normal Case

Suppose the random loss variable *L* evolves as geometric Brownian motion (i.e. a log-normal process) on $[t, T]$ with parameters μ and σ, then the conditional variable $\lambda_{t+\tau} \equiv L_{t+\tau} \mid L_t$ is also a log-normal process. One may equivalently assert that $\mathrm{Ln}(\lambda_{t+\tau})$ is a Gaussian random variable with a mean $\mu\tau + \mathrm{Ln}(L_t)$ and variance $\sigma^2\tau$ which, in standard notation, one denotes as [40, pp. 304-309]:

$$\mathrm{Ln}(\lambda_{t+\tau}) \sim N(\mu\tau + \mathrm{Ln}(L_t), \sigma^2\tau) \tag{B7a}$$

Let Φ denote the cumulative distribution function for the standard normal variable, and let X denote a lognormal random variable with mean *m* and variance s^2; then, according to Dhaene, *et al*. [13], the level-α conditional tail expectation of X is given:

$$\mathrm{CTE}_\alpha(X) = \frac{\Phi(\sigma - \Phi^{-1}(\alpha))}{1-\alpha} \exp(m + 0.5s^2) \equiv \frac{1 - \Phi(\Phi^{-1}(\alpha) - \sigma)}{1-\alpha} \exp(m + 0.5s^2) \tag{B7b}$$

Let z denote the normalised or standard version of the Gaussian random variable described by B7a; and let z_α denote the α-quantile of the said normalised variable, i.e. $\Phi(z_\alpha) = \alpha \iff z_\alpha = \Phi^{-1}(\alpha)$. It follows immediately from B7a and B7b that the level-α conditional tail expectation of the lognormal variable $\lambda_{t+\tau} \equiv L_{t+\tau} \mid L_t$ is:

$$\mathrm{CTE}_\alpha(\lambda_{t+\tau}) = \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1-\alpha} \right] \exp(\mu\tau + \mathrm{Ln}[L_t] + 0.5\tau\sigma^2) \tag{B8a}$$

$$\mathrm{CTE}_\alpha(\lambda_{t+\tau}) = L_t \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1-\alpha} \right] \exp(\tau[\mu + 0.5\sigma^2]) \tag{B8b}$$

Suppose the planning period $[t, T]$ is sub-divided into n intervals of length τ years as shown in Figure B1; let $H_k^T(L_T, \tau)$ denote the iterated CTE at time k of the liability L_T due at time $t = T$; then over the final τ years of the planning period (i.e. at time $t = T - \tau$) we have that:

$$H_{T-\tau}^T(L_T, \tau) \equiv CTE_\alpha(L_T) = L_{T-\tau} \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right] \exp(\tau[\mu + 0.5\sigma^2]) \tag{B9}$$

Proceeding backwards in steps of τ years, given that the CTE is positive homogeneous,[21] we have that:

$$H_{T-2\tau}^T(L_T, \tau) \equiv CTE_\alpha(CTE_\alpha(L_T | L_{T-\tau}) | L_{T-2\tau}) \tag{B10b}$$

$$H_{T-2\tau}^T(L_T, \tau) = CTE_\alpha(L_{T-\tau} | L_{T-2\tau}) \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right] \exp(\tau[\mu + 0.5\sigma^2]) \tag{B10c}$$

$$H_{T-2\tau}^T(L_T, \tau) = L_{T-2\tau} \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right]^2 \exp(2\tau[\mu + 0.5\sigma^2]) \tag{B10d}$$

Similarly, it is easy to show that

$$H_{T-3\tau}^T(L_T, \tau) = L_{T-3\tau} \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right]^3 \exp(3\tau[\mu + 0.5\sigma^2]) \tag{B10e}$$

And so a general formula that applies at any stage k may be derived as follows. Let elements of the set $\mathbb{N}_T \equiv \{1, 2, \cdots, T\}$ denote the number of iteration stages; define a 'stage locator variable' by $m = \eta_i \tau$; then at any decision stage $k \in [t, T] \cap \{m : m = \eta_i \tau, \ \eta_i \in \mathbb{N}_T\}$, we have that $k = T - m$, and the iterated CTE is:

$$H_k^T(L_T, \tau) = L_k \left[\frac{1 - \Phi(z_\alpha - \sigma\sqrt{\tau})}{1 - \alpha} \right]^{\left(\frac{m}{\tau}\right)} \exp(m[\mu + 0.5\sigma^2]) \tag{B11}$$

V. Implementing the Iterated ICTE: A Proposal

In order to use equation (B11), one needs specific values for μ and σ, as well as a tractable model for the random coefficient L_k. Given the requisite data, the former may be obtained by statistical regression analyses; the latter may be modelled as follows. The portfolio's value at time k is given by:

$$v_k = \sum_{i=1}^N p_k^i n_k^i \tag{B12}$$

where: p_k^i is the share price of the i-th asset;

n_k^i is the number of shares held of the i-th asset.

If the risk-free asset bears an interest rate of δ, then the loss variable L_k may be defined as [see **Note** 2 below]:

$$L_k \equiv -Min\{v_k - \bar{v}_k, 0\} \tag{B13a}$$

$$= -Min\{(v_k - w_0 \exp(\delta k)), 0\} \tag{B13b}$$

---

[21] Alternatively, one could also use 'the rule of average conditional expectations' which states [52, pp.402-403]: $\mathbb{E}(Y) = \mathbb{E}[\mathbb{E}(Y | X)]$

In equation (B12), p_k^i is an exogenously driven random variable that is normally modelled as geometric Brownian motion [40, chap. 11]; and n_k^i is a deterministic variable that is a function of the previous investment decisions, i.e., at time k, the proportion of wealth invested in the i-th asset is $u_{k-1}^i v_{k-1}$, and the number of shares held is:

$$n_k^i = \frac{u_{k-1}^i}{p_{k-1}^i} v_{k-1} \qquad (B14)$$

Hence:

$$v_k = v_{k-1} \sum_{i=1}^N \left(\frac{p_k^i}{p_{k-1}^i} \right) u_{k=1}^i \qquad (B15)$$

Assuming no dividends are paid out over the period [0, T], then the i-th asset return at time k is given by:

$$\pi_k^i = \frac{p_k^i - p_{k-1}^i}{p_{k-1}^i} \qquad (B16)$$

Therefore:

$$v_k = v_{k-1} \sum_{i=1}^N \left(\pi_k^i + 1 \right) u_{k-1}^i \qquad (B17)$$

Note that equation (B17) is stochastic due to the presence of the random variables π_k^i. But since these random terms are yet unrealized, they would normally be replaced by numerical estimates $\hat{\pi}_k^i$ computed by some means or other. Granted this, equation (B17) then becomes a tractable first-order difference equation that has the following closed-form solution; see e.g. [8, Chap. 16]:

$$v_k = w_0 \prod_{j=1}^{k-1} \varpi_j \qquad (B18a)$$

with

$$\varpi_j \equiv \sum_{i=1}^N \left(\hat{\pi}_j^i + 1 \right) u_{j-1}^i \qquad (B18b)$$

And hence the coefficient L_k in equation (8) may be modelled as:

$$L_k = -\text{Min}\{\beta_k, 0\} \qquad (B19a)$$

with

$$\beta_k \equiv w_0 \left(1 - \exp(\delta k) \right) \prod_{j=1}^{k-1} \varpi_j \qquad (B19b)$$

Notes on the Proposal

1. In order to use the ICTE risk model just described, one needs numerical estimates for the parameters μ and σ of the assumed probability distribution of the returns, as well as the projected asset returns $\hat{\pi}_k^i$. Note that, unlike the MV approach, this method does not require one to estimate the covariance matrix of the asset returns, and this significantly lowers the amount of data required to run the model.

2. The logic underlying the ICTE risk model is an attempt to encapsulate—at least indirectly—the 'Safety of Principal First' investment strategy and the notion of a 'disaster return' as advocated by Roy [57]. This should be clear on noting that an implicit assumption in the initial definition of L_k in equation (B13) is that, in holding onto the N risky assets, the investor continuously incurs an opportunity cost valued as the gain he foregoes from not having invested all his money into a risk-free asset and holding this up to the current time. Alternatively, one may envisage the initial wealth w_0 as a fixed-term loan at a nominal rate of interest of δ that is due in T years. Either of these interpretations leads to a meaningful definition of downside risk and Roy's 'disaster return' in the portfolio optimization context.

3. The 'min' operator in equations (B13) serves to select only the downside of $(v_k - \bar{v}_k)$, i.e. when the portfolio value is less than the 'disaster value'—the notional opportunity cost or loan liability that accumulates at the rate δ per annum. Equation (B13b) follows from the further assumption that this opportunity cost (or loan liability) is compounded continuously at δ—an investment of w_0 at time $t = 0$ in the risk-free asset held until time k yields $\bar{v}_k = w_0 \exp(\delta k)$.

4. An extended formulation of MP5 that affords more realism may be obtained by augmenting the state vector by an additional coordinate representing the portfolio value V, viz.:

MP12: $\qquad \underset{\bar{\mathbf{u}}}{\mathbf{Opt}} \left\{ g(\bar{\mathbf{u}}_{T-1}), h(\bar{\mathbf{u}}_{T-1}) \right\}$

$$\text{Subject to:} \quad x_{k+1}^j = x_k^j + u_k^j; \qquad x_0^j = 0 \qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\};$$

$$x_{k+1}^{n+1} = x_k^{n+1} \sum_{i=1}^{n} \left(\hat{\pi}_{k+1}^i + 1 \right) u_k^i; \quad x_0^{n+1} = w_0 \qquad k \in \{0, 1, 2, \ldots, T\};$$

$$\sum_{j=1}^{n} u_k^j = 1; \quad u_k^j \in [0, 1] \qquad j \in \{1, 2, \ldots, n\}; \qquad k \in \{0, 1, 2, \ldots, T\};$$

This M-program may be extended even further by explicitly modelling the asset returns $\hat{\pi}_k^i$—as say multi-factor stochastic models in the vein of Glasserman & Xu [24]—and appending the same to the state vector. And depending on whether the formulation is favourable, it may even be possible for well established concepts from control theory (such as the certainty equivalent principle) to be brought to bear.

5. The definition of 'asset returns' in equation (B16) assumes no dividends are paid out in interim period. This is very plausible if one assumes that the planning horizon [0, T] refers to a single financial year, and its sub-divisions τ refer to financial quarters.

6. The 'Balance Sheet Identity' states: Assets – Liabilities = Owner's Equity. At any time k, the value of assets held is v_k; the (notional) liabilities are valued $\bar{v}_k$; and the quantity $v_k - \bar{v}_k$ is therefore a measure of the 'Owner's Equity'. Insolvency occurs when the 'owner's equity' is negative. Thus, as defined above, the ICTE risk measure is an attempt to avoid insolvency over the period [0, T].

7. An empirical validation of the ICTE criterion in the portfolio optimization context is beyond the scope circumscribed for this paper. It requires appropriate data on the probability distribution of the loss variable *L* that allows one to estimate the parameters μ and σ, and until the results of such research are known, this account should not be treated as the definitive approach for incorporating dynamic risk measures in the framework of MP5 or MP12 but merely as a possible way forward. And since the aim here is merely to present alternative approaches to the MV optimization problem and illustrate how GENO may be used to find practical solutions, the risk measure employed in the single-period numerical example to follow shall be 'portfolio variance' as opposed to the more realistic ICTE risk measure.

8. Both models MP5 and MP12 may be run in single-period mode by merely declaring $k \in \{0\}$; and of course the multi-period risk measure *h* may equally be used in the pseudo-dynamic model MP4.

VI. Summary and Conclusions

A risk measure is a function that assigns a real number to the outcome of a random process. This Appendix has presented a measure called the 'Iterated Conditional Tail Expectation' (ICTE) that may be used in a multi-period setting. The ICTE measure exhibits several desirable features: it is 'coherent' in the sense of Artzner, et al. [3, 4]; it is 'consistent' and 'relevant' in the sense of Riedel [55]; and, in most practical cases, it is also 'time-step increasing' in the sense of Hardy & Wirch [26]. Furthermore, in the case where the loss variable is described by a geometric Brownian motion, the ICTE may be expressed in closed form, and this renders it highly practical as a dynamic measure of risk.

References

1. Amilon, H.: Comparison of mean-variance and exact utility maximization in stock portfolio selection. *Scandinavian Working Papers in Economics*, **2001-4**, Department of Economics, Lund University, Sweden, (2001).

2. Artzner, P., Delbaen, F., Eber, J-M., Heath, D., Ku, H.: Coherent multi-period risk adjusted values and Bellman's principle. *Ann. Oper. Res*., **152**, 5-22 (2006).

3. Artzner, P., Delbaen, F., Eber, J-M., Heath, D.: Coherent measures of risk. *Math. Financ.*, **9**, 203-228 (1999).

4. Artzner, P., Delbaen, F., Eber, J-M., Heath, D.: Thinking coherently. *RISK*, **10**, November, 68-71 (1997).

5. Bagamery, B. D.: Ranking differences among the Sharpe ratio and related total portfolio risk adjusted performance Measures. *Journal of Business and Behavioural Sciences*, **1**, 10-21 (2001).

6. Becker, F., Gürtler, M., Hibbeln, M.: Markowitz versus Michaud: Portfolio optimization strategies reconsidered. *Institut für Finanzwirtschaft Working paper*, **IF30V3**, Technische Universität Braunschweig (2010).

7. Bellman, R.: *Dynamic Programming*. Princeton University Press, Princeton NJ (1957).

8. Chiang, A. C.: *Fundamental Methods of Mathematical Economics*. McGraw-Hill, Singapore (1984).

9. Chopra, V. K., Ziemba, W. T.: The effect of errors in means, variances and covariances on optimal portfolio choice. *J. Portfolio Manage.*, **7**, 6-11 (1993).

10. Coello Coello, C. A.: Handling preferences in evolutionary multi-objective optimization: A survey. In: *2000 Congress on Evolutionary Computation*, **1**, 30-37, Piscataway, New Jersey (2000).

11. DeMiguel, V., Garlappi, L., Uppal, R.: Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Studies*, **5**, 1915-1953 (2009).

12. DeMiguel, V., Nogales, F. J.: Portfolio selection with robust estimation. *Oper. Res.*, **57**, 560-577 (2009).

13. Dhaene, J., S. Vanduffel, Q, Tang, M. Goovaerts, R. Kaas, Vyncke, D.: Solvency capital, risk measures and co-monotonicity: A review. *Working paper*, K.U. Leuven, Belgium (2004).

14. Dowd, K.: Adjusting for risk: An improved Sharpe ratio. *Int. Rev. Econ. Financ.*, **9**, 209-222 (2000).

15. Ehrgott, M., Klanroth, K., Schwehm, C.: An MCDM approach to portfolio optimization. *European J. of Oper. Res.*, **155**, 752-770 (2004).

16. Embrechts, P., McNeil, A., Straumann, D.: Correlation and dependency in risk management: properties and pitfalls. *Preprint*, Swiss Federal Institute of Technology, Zürich, Switzerland (1999).

17. Embrechts, P., McNeil, A., Straumann, D.: Correlation: Pitfalls and alternatives. *RISK Magazine*, May, 69-71 (1999).

18. Embrechts, P.: Extreme value theory: Potential and limitations as an integrated risk management tool. *Derivatives Use, Trading & Regulation*, **6**, 449-456 (2000).

19. Fabozzi, F. J., Gupta, F., Markowitz, H. M.: The legacy of modern portfolio theory. *J. Invest.*, Fall, 7-22 (2002).

20. Fabozzi, F. J., Huang, D., Zhou, G.: Robust portfolios: Contributions from Operations Research and Finance. *Ann. Oper. Res*., **176**, 191-220 (2010).

21. Farinelli, S., Ferreira, M., Rossello, D., Thouny, M., Tibiletti, L.: Beyond Sharpe ratio: Optimal asset allocation using different performance ratios. *J. Bank. Financ.*, **10**, 2057-2063 (2008).

22. Freshfields, Bruckhaus, Deringer: German investment modernisation act: Changes to German investment fund and investment fund taxation law (2004). [Online] Available from: http://www.freshfields.com/publications/en.asp, [Accessed March 15, 2007]

23. Fung, W., Hsieh, D. A.: Is mean-variance analysis applicable to hedge funds? *Econom. Lett.*, **62**, 53-58 (1999).

24. Glasserman, P., Xu, X.: Robust portfolio control with stochastic factor dynamics. To appear in *Oper. Res*. (2012).

25. Grossmann, I. E.: Review of nonlinear mixed integer and generalized convex disjunctive programming techniques. *Optim. Eng.*, **3**, 227-252 (2002).

26. Hardy, M. R., Wirch, J. L.: The iterated CTE: A dynamic risk measure. *N. Am. Acturial J.,* **8**, pp 62-75 (2004).

27. Harvey, A.: *The Econometric Analysis of Time Series*. Phillip Allen, London (1990).

28. Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M., Qi, Y., M..: Computing the non-dominated surface in tri-criterion portfolio selection. *Oper. Res*., **61**, 169-183 (2013).

29. Hlawitschka, W.: The empirical nature of Taylor-series approximations to expected utility. *Am. Econ. Rev.*, **84**, 713-719 (1994).

30. Hurlimann, W.: Distortion risk measures and economic capital, *N. Am. Actuarial J.,* **8**, 86-95 (2004).

31. Ishibuchi, H., Tanaka, H.: Multi-objective programming in optimization of the interval objective function. *European J. of Oper. Res.*, **48**, 219-225 (1990).

32. Jiang, C., Han, X., Liu, G. R., Liu, G. P.: A nonlinear interval number programming method for uncertain optimization problems. *European J. of Oper. Res.*, **188**, 1-13 (2008).

33. Kaplan, P. D., Knowles, J. A.: Kappa: A generalized downside risk-adjusted measure. *J. Perform. Measure.*, **8**, 42-54 (2004).

34. Kolm, P.: 60 years of portfolio optimization: Practical challenges and current trends. *Conference Celebrating the 60th Anniversary of Modern Portfolio Theory*. [Online] Available from: www.slideshare.net, November, (2012).

35. Kroll, Y., H. Levy, H., Markowitz, H. M.: Mean-variance versus direct utility maximization. *J. Financ.*, **39**, 47-61 (1984).

36. Kulisch, U.: *Complete Interval Arithmetic and its Implementation on the Computer*. Springer, Berlin (2009).

37. Kydland, F. E., Prescott, E. C.: Rules rather than discretion: The inconsistency of optimal plans. *J. Polit. Econ.*, **85**: 473-491 (1977).

38. Levy, H., Markowitz, H. M.: Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.*, **69**, 308-317 (1979).

39. Luce, R., Raiffa, H.: *Games and Decisions: Introduction and Critical Survey*. Dover Publications, New York (1985).

40. Luenberger, D. G.: *Investment Science*. Oxford University Press, New York (1998).

41. Markowitz, H. M.: Foundations of portfolio theory. *J. Financ.*, **46**, 469-477 (1991).

42. Markowitz, H. M.: Portfolio selection. *J. Financ.*, **7**, 77-91 (1952).

43. Michalewicz, Z., Fogel, D. B.: *How to Solve It. Modern Heuristics*. Springer-Verlag, Berlin (2000).

44. Michaud, R. O, Esch, D. N., Michaud, R. O.: Deconstructing Black-Litterman: How to get the portfolio you already knew you wanted. *J. Invest. Manage.*, **11**, 6-20 (2013).

45. Michaud, R. O.: An Introduction to re-sampled efficiency. *Newsletter*, New Frontier Advisors LLC (2002)

46. Michaud, R. O.: The Markowitz optimization enigma: Is 'optimized' optimal? *Financ. Anal. J.*, **45**, 31-42 (1989).

47. Mina, J., Xiao, J. Y.: Return to RiskMetrics: The evolution of a standard. *Tech. Rep.*, RiskMetrics Group Inc., New York, (2001).

48. More, R. E., Kearfott, R. B. Cloud, M. J.: *Introduction to Interval Analysis*. SIAM, Philadelphia, (2009).

49. Müller, S. M., Machina, M. J.: Moment preferences and polynomial utility. *Econ. Lett.*, **23**, 349-353 (1987).

50. Nawrocki, D.: A brief history of down-side risk. *J. Invest.*, **8**, 9-25 (1999).

51. Niedermayer, A., Niedermayer, D.: Applying Markowitz's critical-line algorithm. In: *Handbook of Portfolio Construction* (pp. 383-400). Springer US (2010).

52. Pitman, J.: *Probability*. Springer-Verlag, New York (1993).

53. Pulley, L. B.: A general mean-variance approximation to expected utility for short holding periods. *J. Financ. Quant. Anal.*, **16**, 361-373 (1981).

54. Rabin, M.: Risk aversion and expected utility theory: A calibration theorem. *Econometrica*, **68**, 1281-1292 (2000)

55. Riedel, F.: Dynamic coherent risk measures. *Stochastic Processes and their Application*, **112**, 185-200 (2004).

56. Rockafellar, R. T.: Coherent approaches to risk in optimization under uncertainty. *INFORMS Tutorials in Operations Research*, Institute for Operations Research and the Management Sciences, Hanover, MD. 38-61 (2007).

57. Roy, A. D.: Safety first and the holding of assets. *Econometrica*, **20**, 431-449 (1952).

58. Salukvadze, M. E.: Optimization of vector functionals. Part I: Programming of optimal trajectories (in Russian). *Avtomatika i Telemekhanika*, **8**, 5-15 (1971).

59. Sharpe, W.: Capital asset prices with and without negative holdings. *J. Financ.*, **46**, 489-509 (1991)

60. Sharpe, W.: Mutual fund performance. *J. Bus.*, **39**, 119-138 (1966).

61. Siwale, I.: GENO™ 2.0: The GAUSS User's Manual. *Tech. Rep. RD-13-2013*, Apex Research Ltd, London (2013). Available with a trial version of GENO—contact Aptech Systems Inc.

62. Siwale, I.: Practical multi-objective programming. *Tech. Rep. RD-14-2013*, Apex Research Ltd, London (2013). [Online] Available at: http://www.researchgate.net or at: http://www.optimization-online.org

63. Siwale, I.: Solution of nonlinear equations via optimization. *Tech. Rep. RD-15-2013*, Apex Research Ltd, London (2013). [Online] Available at: http://www.researchgate.net or at: http://www.optimization-online.org

64. Statman, M.: How many stocks make a diversified portfolio? *J. Financ. Quant. Anal.*, **3**, 353-363 (1987).

65. Stein, M., Branke, J., Schmeck, H.: Portfolio selection: How to integrate complex constraints. *Institute for Applied Informatics and Formal Description Methods. University of Karlsruhe (TH)*, (2005).

66. Steinbach, M. C.: Markowitz revisited: Mean-variance models in financial portfolio analysis. *SIAM Rev.*, **43**, 31-85 (2001).

67. Steuer, R. E., Qi, Y., M. Hirschberger, M.: Developments in multi-attribute portfolio selection. *Multiple Criteria Decision Making '05* (pp. 251-262), (2006).

68. Steuer, R. E., Qi, Y., M. Hirschberger, M.: Suitable-portfolio investors, non-dominated frontier sensitivity, and the effect of multiple objectives on standard portfolio selection. *Ann. Oper. Res*., **152**, 297-317 (2007).

69. Szegö, G.: Measures of risk. *J. Bank. Financ.*, **26**, 1253-1272 (2002).

70. Tütüncü, R. H., Koenig, M.: Robust asset allocation. *Ann. Oper. Res*., **132**, 157-187 (2004).

71. Valente, F., Mitra, G., Poojari, C. A., Kyriakis, T.: Software tools for stochastic programming: A Stochastic Programming Integrated Environment (SPInE): In: *Applications of Stochastic Programming*, SIAM, Philadelphia (2005).

72. Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton NJ (1947).

73. Voorneveld, M., Nouweland, A. V., McLean: Axiomatizations of the Euclidean compromise solution. *Int. J. Game Theory*, **40**, 427-448, (2010).

74. Wang, S. S.: Insurance pricing and increased limits rate making by proportional hazard transforms. *Insurance Math. Econom.*, **17**, 43-45 (1995).

75. Wang, T.: A class of dynamic risk measures. *Working Paper*, University of British Columbia, Canada (2000).

76. Yu, P. L.: A class of solutions for group decision problems. *Manage. Sci*., **19**, 936-946 (1973).

77. Zeleny, M.: Compromise programming. In: *Multiple Criteria Decision Making* (pp.262-301). University of South Carolina Press, Columbia (1973).