Portfolios from Sorts

Robert Almgren[*] and Neil Chriss[†]

April 26, 2005

Abstract

Modern portfolio theory produces an optimal portfolio from estimates of expected returns and a covariance matrix. We present a method for portfolio optimization based on replacing expected returns with *ordering information*, that is, with information about the order of the expected returns. We give a simple and economically rational definition of optimal portfolios that extends Markowitz' mean-variance optimality condition in a natural way; in particular, our construction allows full use of covariance information. We also provide efficient numerical algorithms. The formulation we develop is very general and is easily extended to a variety of cases, for example, where assets are divided into multiple sectors or there are multiple sorting criteria available.

[*]University of Toronto, Departments of Mathematics and Computer Science and Mathematical Finance Program, Robert.Almgren@utoronto.ca

[†]SAC Capital Management, Neil.Chriss@sac.com

THIS PAPER DEVELOPS a method for constructing optimal portfolios from *ordering* information about expected returns. We know of no other systematic methodology for producing optimal portfolios in this context, but we believe that our approach provides a useful addition to the literature because of the vast amount of research that is either directly or indirectly associated with the order of expected returns. Our methods are analogous to mean-variance optimization (Markowitz 1952) in the sense that we use information about both expected returns and risk to produce an optimal portfolio.

To define the meaning of optimal, we make a single economic assumption: investors should prefer to hold portfolios that have higher expected returns in every scenario that is consistent with their beliefs. Mathematically, this amounts to analyzing the set of all expected returns consistent with investor beliefs. The resultant analysis leads to a definition of portfolio optimality and specific computational methods. In this paper we study these methods, present algorithms for computing optimal portfolios and provide empirical evidence for the superior investment performance of these portfolios.

The use of asset sorts in investment theory and practice is well established. A wealth of evidence connects numerical factors associated with specific equities to expected returns. For example, firm characteristics (Fama and French 1992; Banz 1981; Chan and Lakonishok 2004) and prior price history (Asness 1997; Campbell, Grossman, and Wang 1993) have been shown to be related to expected returns. In most cases, the evidence relating numerical factors to expected returns has been offered on the basis of sorting the stocks in the portfolio according to a particular numerical factor and demonstrating a general correlation between the factor values and the average returns. If one believes that this correlation will persist going forward but that the particular functional form of the relationship is difficult to measure, then one might prefer to build optimal portfolios on the basis of the *ordinal* information contained in the factors.

The method in this paper starts with an information set consisting of a covariance matrix and a portfolio sort. A portfolio sort is essentially an estimate of the order of expected returns, from highest to lowest. More generally, one may start with *ordering information* which provides very general information about the relationship between expected returns of the stocks in the portfolio. The heart of the paper is the observation that for a given level of risk an investor should prefer to hold a portfolio that has a higher expected return for *every* expected return consistent with the ordering information and that this yields a preference relation on investment portfolios which allows one to choose among portfolios of equal levels of risk and the portfolio choice problem becomes that of finding the *most preferable* portfolio for a given level of risk. This puts portfolio choice in the context of ordering information on the same intellectual footing as mean-variance optimization, though the details of the analysis are quite different.

The mathematics underlying these methods is more complicated than that of mean-variance optimization, and involves nonsmooth convex optimization and the geometry of conic subsets of $\mathbb{R}^n$ (Boyd and Vandenberghe 2004). In this paper we focus mainly on the economic intuition behind portfolio choice, the implementation and the results. We supply most of the mathematical proofs in a seperate, much longer paper (Almgren and Chriss 2004), which we cite here as AC.

The purpose of this paper is to introduce the notion of optimal portfolios from sorts. We develop the theory in the case of a single complete sort to build intuition and present the theory in with complete clarity. The real strength of this approach, however, is its applicability to general ordering information. The same preference relation construction that leads to optimal portfolios in the case of a single complete sort produces optimal portfolios for a multitude of important practical cases.

The rest of this paper is organized as follows. In Section 1 we give a precise definition of ordering information. In Sections 2–4 we focus on the economic intuition and construction of optimal portfolios for the case of a single complete sort. We do this to make explicit both the construction of the preference relation and the structure of optimal portfolios. In Section 5 we review an empirical study that offers evidence that this method can produce substantial gains over simpler methods. In Section 6 we briefly summarize how our procedure is implemented in practice.

1 Defining sorts

We shall write $S_1, \ldots, S_n$ for the available investment universe of n stocks. In its most general sense, a *portfolio sort* is a set of inequality relationships between the expected returns of these assets. The simplest and most common example is a *single complete sort* which orders all the assets of the portfolio by expected return from greatest to least. If we denote these expected returns as $r_1, \ldots, r_n$, then we mean simply

$$r_1 \geq r_2 \geq \cdots \geq r_n.$$

A sort is a set of beliefs about the first moments of the joint distribution of returns. The underlying assumption is that there is a definite, fixed joint return distribution, whose first moments are the expected returns $r_1, \ldots, r_n$. In this view, available information is not adequate to generate numerical estimates of $r_1, \ldots, r_n$, but it does provide estimates of their order. In practice, sorts may arise from a variety of sources including accounting numbers proxying for measures of value (Chan and Lakonishok 2004) and return histories (Asness 1997; Campbell, Grossman, and Wang 1993). The challenge addressed in this paper is to define what it means for a portfolio to be optimal in relation to ordering information, and then to demonstrate how to calculate optimal portfolios in a variety of circumstances.

A single complete sort may also be expressed as a collection of $n-1$ inequalities

$$r_1 - r_2 \geq 0, \qquad r_2 - r_3 \geq 0, \qquad \cdots, \qquad r_{n-1} - r_n \geq 0.$$

Each of these relations is a *linear homogeneous* inequality, comparing a linear combination of the expected return components to zero. Although our approach was originally motivated by the single complete sort, the theoretical framework that ensues provided optimal portfolios and algorithms for calculating them for a wide range of ordering information that may be useful in practice. We list a few such cases and how they are expressed as ordering information now.

Sector based sorts Our stock universe is divided into k sectors, of size $m_1, \ldots, m_k$ respectively, and we have single complete sorts available *within* each sector, but no information about relative returns *between* sectors; for example, one might have a different sorting methodology for each of the ten S&P GICS sectors of a portfolio of US equities and have no global belief about the order of expected returns.

Deciles and other divisions We divide the stocks into k groups of sizes $m_1, \ldots, m_k$, and we believe that each stock in the first group will outperform each stock in the second group, which will outperform each stock in the third group, *etc*. Such a sorting into deciles has been a mainstay of the research surrounding the relation between accounting based sorts and asset returns (Chan and Lakonishok 2004).

Single complete sort with longs and shorts We have a single complete sort of our n stocks. In addition, we identify ℓ "long names", that is, ℓ assets with positive expected returns, and $n - \ell$ "short names" with negative expected returns. This is the collection of n inequalities

$$r_1 \geq \cdots \geq r_\ell \geq 0 \geq r_{\ell+1} \geq \cdots \geq r_n.$$

Index over/under perform We define an *index* by a linear combination $\mu_1 S_1 + \cdots + \mu_n S_n$, with each $\mu_j > 0$ and $\mu_1 + \cdots + \mu_n = 1$ (for example, the S&P 500). We believe that the first ℓ stocks will *overperform* the index, and the last $n - \ell$ will *underperform:*

$$\begin{aligned} r_j &\geq \mu_1 r_1 + \cdots + \mu_n r_n \quad \text{for } j = 1, \ldots, \ell, \text{ and} \\ r_j &\leq \mu_1 r_1 + \cdots + \mu_n r_n \quad \text{for } j = \ell + 1, \ldots, n. \end{aligned}$$

Higher order sorts A belief frequently expressed in practice is that not only are the assets sorted into decreasing order of expected return, but that the *spreads* between expected returns are greater for certain pairs than for others. For example, one might believe not only that $r_1 \geq \cdots \geq r_n$, but also that the spread between the top two names is bigger than the spread between the next two:

$$r_1 - r_2 \geq r_2 - r_3.$$

In a typical application, one might believe that the information is more reliable in the tails, that is, the spreads between returns are steadily decreasing for the first half of the asset list, and steadily decreasing for the second half. This would give approximately $2n$ inequalities. Of course, this construction can be extended to higher order differences.

Multiple sorts Managers will sometimes wish to combine information from multiple indicators. This might arise, for example, in the case where one looks at value and momentum strategies (Asness 1997). Each of these gives a complete or partial sort of some or all of the assets in the universe, but generally the sorts are *incompatible:* only highly degenerate return vectors with many components equal can simultaneously satisfy all the beliefs. Our construction shows how to form an optimal portfolio taking account of all these conflicting beliefs.

We have listed the above possible information structures to give the reader a flavor of the general nature of our formulation and the level of its applicability. In the rest of the paper, for the sake of concreteness we shall focus on the case of a single complete sort. But the techniques and construction presented here extend naturally to all of the above scenarios.

2 Portfolio preference relations

In this section we provide a principle for comparing the preferability of investment portfolios given ordering information about their expected returns, toward the goal of producing optimal portfolios. Stepping back a moment, in a mean-variance optimization framework a portfolio is *efficient* if it delivers the highest level of expected return for a given level of risk. This is the cornerstone of modern portfolio theory, but it depends on having a numerical value for the expected return of each stock in the portfolio. To study portfolio optimization based on ordering information, we must be able to compare portfolios in a manner that does not require direct numerical comparisons. We start by re-examining Markowitz portfolio optimization from the point of view of preference relations.

We may state Markowitz' basic portfolio choice framework in terms of a preference relation. Given a choice between two portfolios delivering equal levels of risk, an investor should choose the one with the higher level of expected return. From this point of view, mean-variance efficient portfolios may be defined as those that are *maximally preferable* among all portfolios with a given level of risk. In this way we move the question of portfolio optimality from numerical computation of maximal expected return to one of maximal preferability, which requires only binary comparison between pairs of candidate portfolios.

We represent a portfolio by a vector $w = (w_1, \ldots, w_n)$, where the components w_j are real numbers representing *dollar investments* in the respective stocks, and any uninvested amount is held in cash. Such investments may be positive (representing long positions) or negative (representing short positions). Thus for an expected return vector $r = (r_1, \ldots, r_n)$, the expected dollar return of the portfolio return is $w \cdot r = w_1 r_1 + \cdots + w_n r_n$, and its dollar variance is $w \cdot V \cdot w$ where V is the $n \times n$ covariance matrix. (We shall leave off the traditional transpose notation: "make the dimensions match".)

In a world where an investor does not know the expected returns of individual stocks he cannot know the expected returns of a portfolio formed from those stocks. We would nevertheless like to be able to place a preference relation on portfolios, stating when an investor should prefer one portfolio to another. To do this we start by observing that some expected return vectors $r = (r_1, \ldots, r_n)$ are *consistent* with given ordering information and some are not. For example, for a single complete sort, consistent expected returns are precisely those vectors r for which $r_1 \geq r_2 \geq \cdots \geq r_n$.

We write Q for the set of all consistent expected returns related to a sort. This set is a key object of study in the construction of optimal portfolios and differentiates our approach from other potential methods of producing optimal portfolios from sorts by in effect studying *all* possible expected returns at once and not playing favorites by choosing one set over another. The set Q is a mathematical space called a cone and we exploit its rich geometric structure to find optimal portfolios (see Boyd and Vandenberghe (2004) for more on the mathematics of cones). We use the set Q in conjunction with the following assumption.

Economic Assumption 1 If w and v are portfolios, then, leaving aside risk limits or other investment constraints, an investor should prefer to hold w over v if the expected return of w is greater than or equal to that of v for *every* consistent expected return vector r, that is, for every $r \in Q$. We write $w \succeq v$ to mean *w is preferable to v*.

This is not a difficult assumption to accept, for it must be true if one believes the basic assumption of mean-variance analysis that an investor will prefer to hold portfolios of higher rather than lower expected returns. Since we believe there is a concrete expected return vector $r = (r_1, \ldots, r_n)$ which is the *actual* expected return vector, we know that if $w \succeq v$ then in particular $w \cdot \rho \geq v \cdot \rho$ since $\rho \in Q$. Therefore one should prefer to hold w to v. In fact, this definition turns out to be a bit stronger than we need to produce optimal portfolios. We see this through the following definition.

For any sorted list of assets there is a set of *fundamental* portfolios with the property that each has a non-negative expected return for *any* expected return $r \in Q$. For a single complete sort, consider first the portfolio $e_1 = (1, -1, 0, \ldots, 0)$, that buys one dollar of S_1 and sells one dollar of S_2. This has a non-negative expected return for any $r \in Q$ since $w \cdot r = r_1 - r_2 \geq 0$. This first fundamental portfolio e_1 is therefore an investment portfolio expressing the belief that S_1 has a higher expected return than S_2. For a complete sort of n stocks, there are $n-1$ such fundamental portfolios. We write e_i for the portfolio that buys one dollar of the i-th stock and sells one dollar of the $(i+1)$-st:

$$e_i = (0, \ldots, 0, 1, -1, 0, \ldots, 0).$$

Each such portfolio expresses a single belief about the expected returns. Conditional on the belief that the sort properly expresses the order of the expected returns, each e_i has a non-negative expected return.

Now, if $\lambda_1, \ldots, \lambda_{n-1}$ are any nonnegative numbers then the portfolio

$$w_\lambda = \sum_{i=1}^{n-1} \lambda_i e_i$$

also has a non-negative expected return for any consistent return vector. Even though we do not know the expected returns of the stocks in the portfolio, and we cannot estimate the expected return of w_λ, we know that it has a non-negative expected return.

Having constructed w_λ we can now show how to produce a portfolio w_μ that is preferable to w_λ. Write

$$w_\mu = \sum_{i=1}^{n-1} \mu_i e_i,$$

and suppose that $\mu_i \geq \lambda_i$ for all i. Then w_μ has an expected return that is bigger than or equal to w_λ, since the *difference* portfolio

$$w_\mu - w_\lambda = \sum_{i=1}^{n-1} (\mu_i - \lambda_i) e_i$$

has weights $\mu_i - \lambda_i \geq 0$ and thus represents a long-only combination of the fundamental portfolios. Therefore on the basis of expected return we clearly prefer portfolio w_μ to w_λ.

The above observation leads to a cleaner definition of a portfolio preference relation definition. First note the portfolio $\mathbb{1} = (1, \ldots, 1)$, which represents a long position in all assets simultaneously; the interesting thing about this portfolio is that the belief it represents is *orthogonal* to that of the sort. We discuss this in more detail below, but for now, given the introduction of $\mathbb{1}$ we note that the collection $(e_1, \ldots, e_{n-1}, \mathbb{1})$ is a complete basis for the n-dimensional space of portfolios. That is, we can write any portfolio w as a sum

$$w \;=\; \sum_i \lambda_i e_i \;+\; \gamma\,\mathbb{1}.$$

Definition of the portfolio preference relation For any portfolios w and v, write

$$w - v \;=\; \sum_{i=1}^{n-1} \lambda_i e_i \;+\; \gamma\,\mathbb{1}.$$

Then $w \succeq v$ if each $\lambda_i \geq 0$ for every i. We ignore the sign of γ.

That is, we prefer w to v if the part of the difference portfolio $w - v$ belonging to the $n - 1$ fundamental portfolios e_i has a positive expected return for every expected return $r \in Q$, that is, for every expected return consistent with the portfolio sort. This is almost the same as saying we prefer w to v if w has a greater expected return than v for every consistent return vector, but not quite the same because of the portfolio $\mathbb{1}$.

As noted above, investments in the fundamental portfolios express beliefs related to the sort. An long investment in $\mathbb{1}$ expresses the belief that the average return of the stocks in the portfolio will be positive, but this average return is completely independent from the sort itself. Indeed, for each consistent expected return vector r with $\mathbb{1} \cdot r \geq 0$, there is another consistent expected return vector $\bar{r}$ with $\mathbb{1} \cdot \bar{r} \leq 0$; $\bar{r}$ is obtained from r by reversing the order of the components and changing their sign (for example, $r = (3, 2, 1) \mapsto \bar{r} = (-1, -2, -3)$). Thus, requiring the difference portfolio to have nonnegative return for every $r \in Q$ would require $\gamma = 0$. The correct portfolio preference relation ignores this irrelevant investment in the mean by ignoring the component along $\mathbb{1}$.

Having defined $w \succeq v$ as above, one may ask what about cases where not all $\lambda_i \geq 0$. If the λ_i have mixed sign then this definition says nothing about the relative preference of w and v. In other words the preference relation suffers from the weakness that it cannot definitively compare all pairs of portfolios. This turns out to be easily fixed by refining the definition of the preference relation, and this is the topic of Section 4.

In the general formulation of our preference relation (see AC), the belief structure given by the ordering information creates an orthogonal decomposition of the space of portfolios into a *relevant* subspace R and an *irrelevant* subspace $R^\perp$. There may or may not exist a convenient set of basis vectors for these subspaces, but any portfolio weight vector w (or difference portfolio) may always be decomposed as

$$w = w_{\mathrm{rel}} + w_{\mathrm{irrel}}, \qquad w_{\mathrm{rel}} \in R, \quad w_{\mathrm{irrel}} \in R^\perp,$$

and we compare portfolios only by comparing the relevant parts w_{rel}.

The same decomposition is implicitly present in mean-variance analysis. Namely, if r is an expected return estimate than one may write

$$w = w_{\mathrm{rel}} + w_{\mathrm{irrel}}, \qquad w_{\mathrm{rel}} = \lambda r, \quad w_{\mathrm{irrel}} \cdot r = 0.$$

For determining the expected return of the portfolio, w_{rel} is the only part that matters (and this is one dimensional!). Of course the irrelevant direction affects the level of risk in the portfolio and must be considered in determining the actual optimal investment, but it is orthogonal to the expected return estimate so has no impact on the portfolio's expected return.

3 Efficient portfolios

In practice, investment portfolios are determined by the interplay between expected return and risk. In our formulation, we look for portfolios that are maximally preferable within given investment constraints that may be based on risk limits or other factors. The preference relation and the constraints are equally important.

Let $\mathcal{M}$ be the set of all portfolios meeting a set of budget constraints. For example, in the Markowitz framework, $\mathcal{M}$ is the set of portfolios with variance less than or equal to a given fixed level. A portfolio w is *efficient* with respect to the budget set $\mathcal{M}$ and the sort if there is no portfolio v also in $\mathcal{M}$ that is strictly preferable to w. That is, there does not exist $v \in \mathcal{M}$ with $v \succeq w$ but $w \not\succeq v$. Conversely, if w is not efficient, then there is an efficient portfolio $v \in \mathcal{M}$ whose relevant component has a higher expected return for *every* expected return consistent with the sort.

This theoretical definition of efficiency is precisely analogous to the definition of portfolio efficiency in mean-variance analysis. We posit an economically motivated preference relation on portfolios and then state that for a particular budget constraint an investor should prefer to hold portfolios which are most preferable with respect to this preference relation. With this stated, the most obvious next challenge is to calculate efficient portfolios and study their properties.

In AC we prove a pair of general theorems that describe efficient portfolios in terms of a relationship between normal vectors to supporting hyperplanes of $\mathcal{M}$ and the set of consistent return vectors. Here we shall characterize efficient portfolios very explicitly for the case of a single complete sort.

We assume we have a single complete sort of the assets $S_1, \ldots, S_n$, and that we have the covariance matrix V. The budget set $\mathcal{M}$ is the set of all portfolios whose risk is less than a fixed level σ^2:

$$\mathcal{M} = \left\{ w \in \mathbb{R}^n \,\middle|\, w \cdot V \cdot w \leq \sigma^2 \right\}$$

Classic mean-variance optimization takes as input a specific expected return vector r and gives the optimal portfolio $w \sim V^{-1} r$, where $\sim$ means "scaled as necessary by a positive factor so that $w \cdot V \cdot w = \sigma^2$."

Our goal now is to show that the set of efficient portfolios for a single complete sort is exactly those that are mean-variance optimal for expected returns that are both consistent with the sort and sum to zero. That is, there is a one-to-one correspondence between efficient portfolios and vectors r such that

$$r_1 \geq r_2 \geq \cdots \geq r_n \quad \text{and} \quad r_1 + \cdots + r_n = 0.$$

To see this, let $E_1, \ldots, E_{n-1}$ be the collection of vectors

$$E_j \;=\; \frac{1}{n}\big(\underbrace{n-j, \ldots, n-j}_{j}, \underbrace{-j, \ldots, -j}_{n-j} \big), \qquad j = 1, \ldots, n-1,$$

so that $E_i \cdot e_j = \delta_{ij}$ and $E_i \cdot \mathbb{1} = 0$; thus E_i represents unit exposure to the ith difference portfolio. Now let $x = (x_1, \ldots, x_{n-1})$ be any vector of numbers such that each $x_i \geq 0$. We prove in AC that efficient portfolios are precisely those that can be written

$$w \sim V^{-1}\big(x_1 E_1 + \cdots + x_{n-1} E_{n-1}\big) \quad \text{with} \quad x_1, \ldots, x_{n-1} \geq 0. \tag{1}$$

And the vector $y = x_1 E_1 + \cdots + x_n E_n$ has decreasing coefficients, because $y_i - y_{i+1} = e_i \cdot y = x_i \geq 0$. The converse statement is not hard to show as well, but we refer the reader to AC for details.

The above characterization of efficient portfolios in terms of mean-variance optimal portfolios is in some sense misleading. In the first place, given a covariance matrix V and a portfolio w we have

$$w = V^{-1} r, \quad r = V w.$$

That is, every portfolio is mean-variance optimal for some expected return vector.

The fact that efficient portfolios are classified specifically as those which are mean-variance optimal for expected returns that are decreasing with the sort and sum to zero is not surprising. What should be surprising, however, is that if a portfolio is not efficient, then first $r = Vw$ is either not consistent with the sort or does not sum to zero and, more importantly, there exists a portfolio v whose relevant part has a higher expected return than w's for *every* expected return consistent with the sort.

Characterization of efficient portfolios Let w be a portfolio with $wVw = \sigma^2$ formed to invest optimally in a portfolio of assets sorted into a single complete sort $S_1, \ldots, S_n$, and $r = Vw$. Suppose that $r = (r_1, \ldots, r_n)$ and either $\sum r_i \neq 0$ or r is not consistent with the sort. Then there is a portfolio v with $vVv = \sigma^2$ such that

$$v - w = \sum \lambda_i e_i + \gamma \mathbb{1}, \quad \lambda_i \geq 0 \text{ for all } i.$$

In other words, if Vw does not meet the consistency test, then there exists a v with the same risk level but with strictly greater exposure to the fundamental portfolio.

The space of efficient portfolios is still quite large as it encompasses essentially every decreasing sequence of returns that sum to zero. Can we distill out of the set of efficient portfolios a single optimal portfolio to trade? The answer is yes and we explain the details in the next section.

4 Optimal portfolios

In this section we show how to choose a single optimal portfolio from the set of efficient portfolios. To do this we clarify the information content implicit in a single complete sort and more generally in ordering information. If we have a portfolio of stocks $S_1, \ldots, S_n$ ordered so that $r_1 \geq \cdots \geq r_n$ then we are positing two things. First, the obvious, that the expected returns of the stocks, or more precisely, the joint distribution of the stocks, respect the ordering. Second, this information is the *only* information we have about the expected returns. If we have more information about the expected returns than the order, we must add this to our ordering information as described in Section 1.

To begin this analysis, we start with a simple observation concerning mean-variance analysis. If r is a vector of expected return estimates for a portfolio of stocks, then the optimal portfolio associated to r does not depend on the magnitude of r once a budget is established. That is, if we repace r with λr, where $\lambda > 0$ is a scalar, then λr and r produce the same optimal portfolios for a given budget constraint; this is a consequence of the scaling for constant risk level.

What this means is that the information in an expected return vector relevant for optimization is contained completely in its vector direction, not in its magnitude. It turns out that this is equally true in the case of ordering information and the portfolio choice theory we build from it, meaning that a given *consistent* expected return vector is only important up to modifications by a positive scalar factor. Therefore in what follows we will refer to *consistent expected return directions* when we wish to focus only on an expected return up to a positive scalar. With this as background we make our key modeling assumption.

Modeling assumption For a single complete sort we assume that each expected return direction is equally likely: there is no bias toward some directions over others. The only information in the model is the sort itself. We express this mathematically by introducing a *radially symmetric* probability measure μ on the space Q of consistent expected returns. That is, μ has the same form on each ray of possible expected returns (a "ray" is the set $\{\lambda r | \lambda \geq 0\}$). The exact form of the measure μ does not matter because optimal portfolios are only determined by the direction, and not the magnitude, of expected returns. μ assigns equal probability to every direction in the space of expected returns.

To define optimal portfolios, we start by recalling that for given portfolios w and v we are only concerned about the *relevant* parts of the portfolios, that is, those coming from the fundamental portfolios which express one of our beliefs about the ordering. Recalling this we may restate the preference relation of section 2 in terms of μ:

$$w \succeq v \quad \text{if and only if} \quad \mu\Big(\big\{\, r \in Q \,\big|\, w_{\text{rel}} \cdot r \geq v_{\text{rel}} \cdot r \,\big\}\Big) = 1,$$

or equivalently,

$$w \succeq v \quad \text{if and only if} \quad \mu\Big(\big\{\, r \in Q \,\big|\, v_{\text{rel}} \cdot r \geq w_{\text{rel}} \cdot r \,\big\}\Big) = 0.$$

In words this says that an investor prefers portfolio w to portfolio v if for 100% of the instances portfolio w has a higher expected return than portfolio v. Posed in this way, it is natural to refine the definition by considering measures between zero and one. This leads us to

Economic Assumption 2 If w and v are portfolio weight vectors, then, leaving aside risk limits or other investment constraints, an investor should prefer w over v if the expected return of w is greater than or equal to that of v for *a greater fraction* of possible expected return vectors.

We thus now refine the preference relation $\succeq$ by defining

$$w \succeq v \quad \text{if and only if}$$
$$\mu\Big(\big\{\, r \in Q \,\big|\, w_{\text{rel}} \cdot r \geq v_{\text{rel}} \cdot r \,\big\}\Big) \geq \mu\Big(\big\{\, r \in Q \,\big|\, v_{\text{rel}} \cdot r \geq w_{\text{rel}} \cdot r \,\big\}\Big).$$

This is a continuous version of $\succeq$. It is an obviously weaker requirement and therefore allows us to compare efficient portfolios.

We now define a portfolio to be optimal with respect to a sort if it is *most preferable* under the preference relation for a given level of risk. This definition is extremely simple and straightforward, but yields no obvious method for calculating optimal portfolios. It turns out the mathematical derivations involve a fairly subtle analysis (the bad news), but the resultant portfolios are extremely easy to calculate (the good news).

To calculate optimal portfolios requires us to characterize $\succeq$ in terms of something concrete. In AC we do this and show that there is a vector c, defined as the center of mass of the set Q, with the following amazing property. If w and v are arbitrary portfolios, then

$$w \succeq v \quad \text{if and only if} \quad w \cdot c \geq v \cdot c.$$

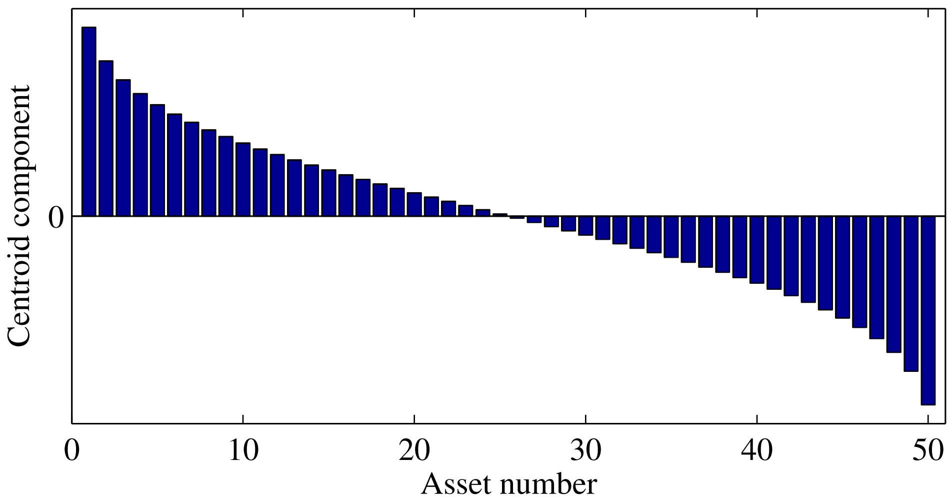

Figure 1: The centroid vector for a complete sort of 50 assets. Relative to a linear profile, the centroid overweights very high and very low ranked stocks while underweighting the middle. Vertical scale is arbitrary.

This means, in particular, that the preference relation is entirely characterized by a simple linear function. This means to find the most preferable portfolio relative to a maximum risk constraint is equivalent to finding the maximum of the linear function c on the set $\mathcal{M}$ of portfolios respecting this constraint. In other words, to find the optimum portfolio we solve the following linear program with quadratic constraints:

$$\max_{w}\, w \cdot c \quad \text{subject to} \quad w \cdot V \cdot w \le \sigma^2$$

where again c is the centroid vector. In AC we call the solution to this problem the *centroid optimal* portfolio. Within the set of efficient portfolios, the centroid is a naturally chosen best element. There is a centroid optimal portfolio for any type of sort as well as multiple sorts. Figure 1 shows the centroid vector c for a single complete sort of $n = 50$ assets.

The centroid vector may be computed by Monte Carlo methods or by evaluation of integrals. In simple cases, analytical approximations can be derived. For example, for a single complete sort of n assets, the jth component of c is approximated to within one-half of one percent by

$$c_{j,n} = N^{-1}\left(\frac{n+1-j-\alpha}{n-2\alpha+1}\right), \quad \alpha = A - Bn^{-\beta},$$

where $N^{-1}(\cdot)$ is the inverse cumulative normal distribution and $A = 0.4424$, $B = 0.1185$ and $\beta = 0.21$. This construction is somewhat reminiscent of "normal scores," but we provide a precise characterization of the offset α as well as a framework that extends to more general scenarios.

Figure 2 shows the centroid vector for several more complicated information structures. In the sector case (top panel), our construction naturally fixes the relative sizes of the extreme values in the two sectors. The decile centroid (bottom panel) is not the same as a centroid of 10 assets.

5 Empirical tests

Theoretical elegance alone does not necessarily deliver improved investment performance. Does our solution actually deliver better performance in practice? And is it robust enough to give good performance in the presence of the ranking errors that are inevitable in real situations? We answer these questions with two series of empirical tests:

- We use **historical** returns data from the CRSP data set to implement the centroid strategy exactly as it would be done in practice. As our forecaster of future expected returns, we use a reversal strategy based on one-week preceding returns; this simplifies the study by eliminating the need for other economic inputs. The centroid optimal strategy delivers substantially higher returns for the same level of risk than do any of the alternatives in current use.

- We generate **simulated** data using specified values for the expected return and their variance-covariance matrix. Thus the order is precisely known, but we introduce random permutations to degrade the quality of the information. The performance of the centroid optimal portfolio decreases only slowly as the severity of the perturbations is increased.

Here we only summarize the results.

The foundation for our historical study is summarized by Thorp (2003):

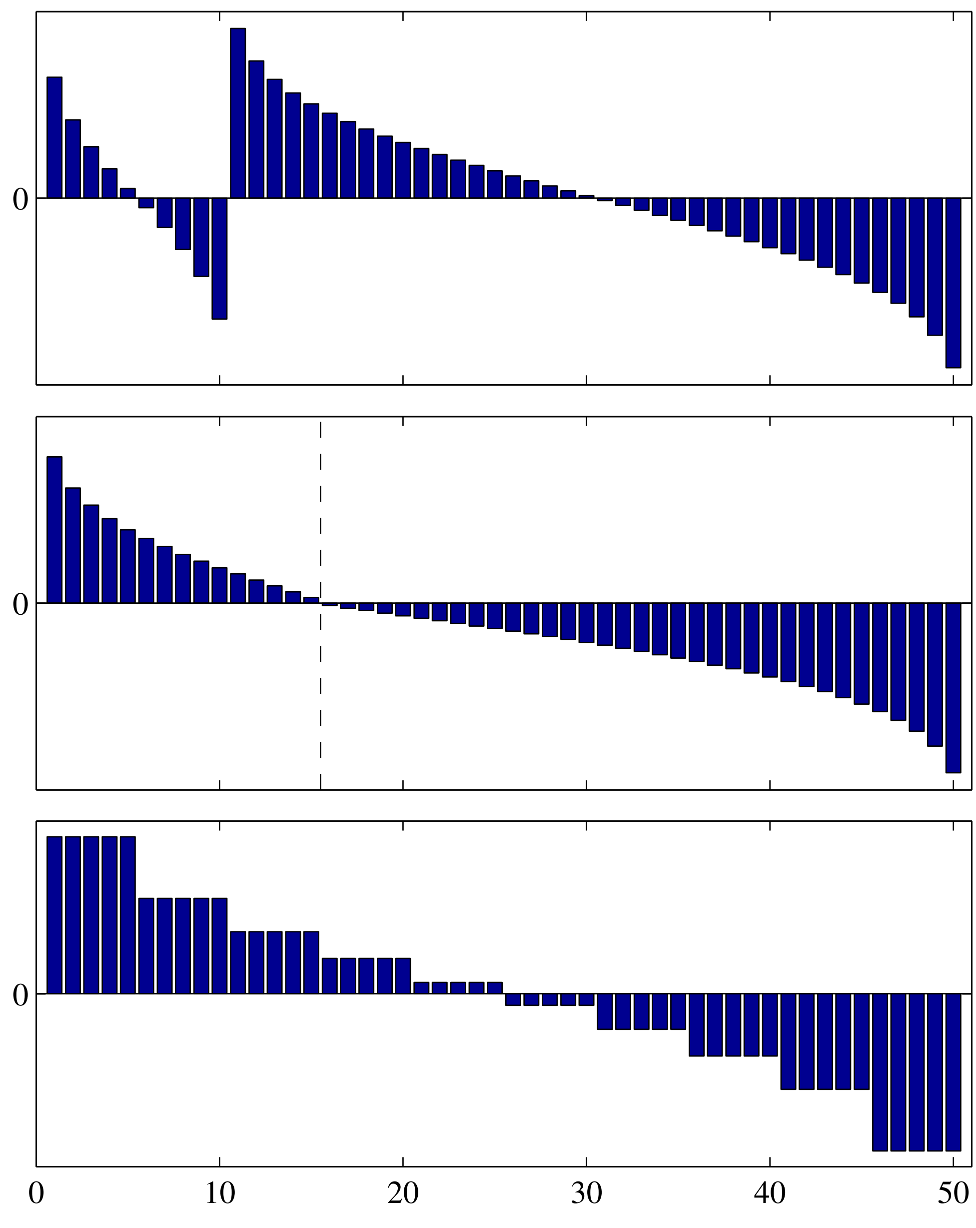

Figure 2: The centroid vector for 50 assets, with different information structures. Top panel: sectors of sizes 10 and 40. Middle: complete sort, with the first 15 assets expected to have positive returns, the last 35 negative. Bottom: 10 deciles of 5 assets each.

> An empirical tendency for common stocks to have short-term price reversal ... was discovered in December 1979 or January 1980 in our shop [Princeton Newport Partners] as part of a newly initiated search for "indicators", technical or fundamental variables which seemed to affect the returns on common stocks. Sorting stocks from "most up" to "most down" by short-term returns into deciles led to 20% annual returns before commissions and market impact costs on a portfolio that went long the "most down" stocks and short the "most up" stocks.

This "reversal effect" was extensively studied and confirmed by Campbell, Grossman, and Wang (1993) (we neglect the role of volume).

From the CRSP database of US stock prices from NYSE, Amex, and Nasdaq, we form an initial portfolio of the 1,000 largest capitalization stocks on January 19, 1990, for which there exist at least 1,000 preceding days of returns data (in order to estimate covariances). We follow this list of stocks day-by-day to December 31, 2002. If a stock disappears from the universe, we replace it by the largest new stock that has at least 1,000 preceding days of data. In this way we form a universe of approximately 3,000 daily returns, with about 2,000 different stocks such that on each date we have at least 1,000 stocks for each of which we have 1,000 preceding days of data. This procedure is free from look-ahead bias.

We study portfolios of varying sizes n. On each day t, we form an estimated covariance matrix from the preceding 1000 days of data. Using the reversal strategy, at each date t we sort the stocks into *increasing* order by their return over the days $t-L-K$ to $t-L$: stocks which decreased the most are expected to increase the most in the next period. Here K is the *reversal period*, and L is the *lag*, representing the delay in implementing the new strategy. Here we describe the results for a reversal period of five days and for lags of zero (the new portfolio is implemented immediately at close using that day's returns) and one day (the new portfolio is implemented the following day). Portfolios are rebalanced daily.

On each date t we form four different portfolios (Table 1):

1. The *unoptimized linear* portfolio takes weights $w \sim \ell$, where ℓ is a linear vector that is long the highest-ranked assets and short the lowest, with linear interpolation; no use is made of covariance information except for the risk scaling described below. This roughly corresponds to the approach described by Thorp.

	Linear	Centroid
Unoptimized	ℓ	c
Optimized	$V^{-1}\ell$	$V^{-1}c$

Table 1: The four portfolios compared in the empirical test. The layout matches the results shown in Table 2.

2. The *unoptimized centroid* portfolio takes the weights to be the centroid vector c rather than linear, giving slightly more weight to the ends.

3. The *optimized linear* portfolio uses the linear vector ℓ as imputed returns rather than weights; the portfolio weights are $w \sim V^{-1}\ell$.

4. The *optimized centroid* portfolio is the one we advocate, using the centroid vector c as imputed returns; the weights are $w \sim V^{-1}c$. By the arguments of this paper, this portfolio should give, on average, the highest return for a given level of risk.

In each case the weight vector is scaled to have unit *ex ante* risk according to the estimated covariance matrix V.

Figure 3 and Table 2 show the results, across a range of portfolio sizes (note that since portfolio volatility has been scaled to one, information ratio is equivalent to average return). Three points are immediately evident:

- Incorporation of covariance information dramatically improves risk-adjusted return (upper two curves). Without the approach outlined in this paper, it is not at all obvious how to combine covariance data with inequality information.

- The centroid vector gives a substantial improvement over the linear vector, even without covariance information. We know of no way to derive the centroid vector without our analysis.

- The centroid portfolio is even better for large portfolios.

Our construction is dramatically better than *ad hoc* alternatives.

Number of stocks	Reversal Period (days) 5		10		15		20		25	
25	2.50	2.47	2.36	2.40	1.72	1.75	1.42	1.45	1.59	1.61
	3.21	3.20	2.37	2.50	1.84	1.95	1.69	1.93	1.63	1.79
50	2.88	2.95	2.92	3.10	2.39	2.52	2.07	2.16	2.06	2.14
	3.53	3.99	3.26	3.63	3.03	3.35	2.93	3.30	2.79	3.07
100	3.18	3.20	2.98	3.12	2.46	2.61	2.09	2.17	2.17	2.19
	4.26	4.76	3.65	4.09	3.54	3.95	3.19	3.73	3.10	3.43
200	3.04	3.20	2.64	2.83	2.40	2.54	2.05	2.20	2.17	2.27
	4.96	5.87	3.81	4.61	3.83	4.51	3.37	4.18	3.08	3.75
500	2.97	3.22	2.40	2.72	2.11	2.37	1.91	2.19	1.93	2.16
	5.82	6.88	4.33	5.38	4.31	5.25	4.44	5.40	4.31	5.10

25	2.32	2.32	2.10	2.15	1.61	1.61	1.20	1.24	1.46	1.46
	2.39	2.41	1.84	1.86	1.35	1.40	1.25	1.43	1.35	1.48
50	2.91	2.97	2.80	2.96	2.29	2.37	1.85	1.90	1.93	1.97
	2.58	2.89	2.52	2.85	2.46	2.65	2.32	2.54	2.30	2.48
100	3.25	3.15	2.77	2.84	2.29	2.38	1.81	1.85	1.95	1.94
	3.07	3.30	2.47	2.83	2.62	3.02	2.34	2.87	2.38	2.68
200	3.25	3.22	2.41	2.49	2.11	2.18	1.67	1.76	1.80	1.89
	3.70	4.13	2.70	3.17	2.73	3.27	2.32	3.01	2.16	2.70
500	2.72	2.84	1.95	2.15	1.60	1.81	1.36	1.59	1.37	1.60
	3.55	4.20	2.27	2.95	2.47	3.19	2.74	3.46	2.71	3.30

Table 2: Information ratios for the four strategies considered in this paper, for varying lag, reversal period, and portfolio size. Upper box is lag of zero days; lower box is lag of one day. Within each box, the layout is as in Table 1: the left column is based on the linear portfolio, the right on the centroid; the upper row is the unoptimized portfolios and the lower row is the optimized portfolios.

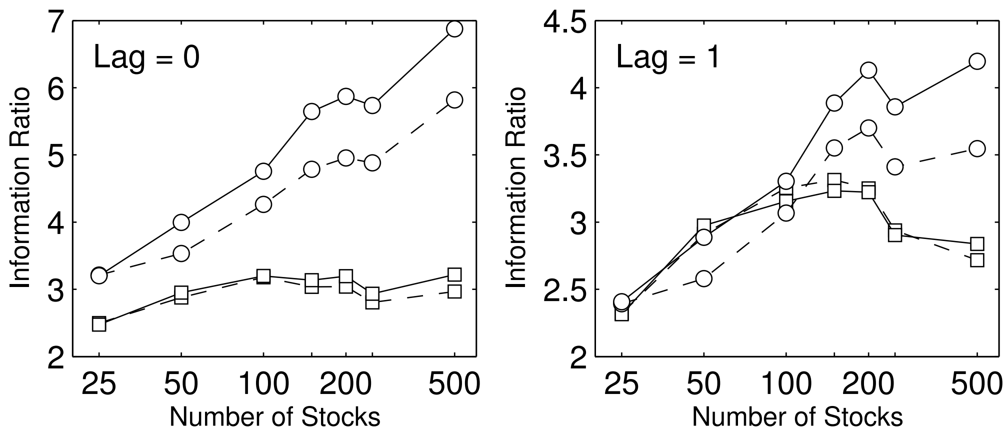

Figure 3: Realized returns for four different algorithms. From bottom to top (it is clearer in the left panel) the curves show linear, centroid, optimized linear, and optimized centroid. For large portfolios, the optimized centroid gives more than a two-fold improvement over the unoptimized linear, and is substantially better than the optimized linear.

6 Summary

We close with a concise summary of our procedure.

1. Identify the set $\mathcal{M}$ of allowable portfolio vectors w, in terms of constraints such as limits on total portfolio risk, total dollar investment, position limits on individual assets, turnover limits, *etc.* Covariance information is included here if desired.

2. List all available information about the expected return vector r, in the form of homogeneous inequality relationships (Section 1). This defines the *consistent cone* Q, the "coarse" preference relation (Section 2), and the set of *efficient* portfolios (Section 3).

3. Compute the centroid vector c of the cone Q. For simple information structures this may be done nearly analytically; for more complicated belief sets it may be computed by Monte Carlo calculation as discussed in AC. The centroid vector defines the "fine" version of the portfolio preference relation (Section 4).

4. Compute the optimal portfolio w by maximising the scalar quantity $c \cdot w$ over $w \in \mathcal{M}$. This is a standard linear programming problem, although the constraints may be complicated. An extremely rich set of tools are available to solve this problem efficiently.

References

Almgren, R. and N. Chriss (2004). Optimal portfolios from ordering information. `http://ssrn.com/abstract=633801`. 3

Asness, C. S. (1997, Mar/Apr). The interaction of value and momentum strategies. *Financial Analysts Journal 53*(2), 29–36. 2, 4, 5

Banz, R. W. (1981). The relationship between return and market value of common stocks. *J. Financial Econ. 9*, 3–18. 2

Boyd, S. and L. Vandenberghe (2004). *Convex Optimization*. Cambridge University Press. 3, 7

Campbell, J., S. J. Grossman, and J. Wang (1993). Trading volume and serial correlation in stock returns. *Quart. J. Econ 108*, 905–939. 2, 4, 18

Chan, L. K. C. and J. Lakonishok (2004, Jan/Feb). Value and growth investing: Review and update. *Financial Analysts Journal 60*(1), 71–86. 2, 4, 5

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *J. Finance 47*, 427–465. 2

Markowitz, H. (1952). Portfolio selection. *J. Finance 7*, 77–91. 2

Thorp, E. (2003). A perspective on quantitative finance: Models for beating the market. In *The Best of Wilmott 1: Incorporating the Quantitative Finance Review*, pp. 33–38. Wiley. 16