Robust Portfolio Optimization Using a Simple Factor Model

Chris Bemis, Xueying Hu, Weihua Lin, Somayeh Moazeni,
Li Wang, Ting Wang, Jingyan Zhang

Abstract

In this paper we examine the performance of a traditional mean-variance optimized portfolio, where the objective function is the Sharpe ratio. We show results of constructing such portfolios using global index data, and provide a test for robustness of input parameters. We continue by formulating a robust counterpart based on a linear factor model presented here. Using a dynamic universe of stocks, the results for this robust portfolio are contrasted with a naive, evenly weighted portfolio as well as with two traditional Sharpe optimized portfolios. We find compelling evidence that the robust formulation provides significant risk protection as well as the ability to provide risk adjusted returns superior to the traditional method. We also find that the naive portfolio we present outperforms the traditional Sharpe portfolios in terms of many summary metrics.

1   Introduction

We begin by examining the performance of a traditional Sharpe ratio optimization problem using global indexes in Section 2. Here we establish the quadratic programming problem of interest to us and exhibit the nonstationarity of the input parameters. We proceed in Section 3 to add additional constraints to the original problem to mimic realistic portfolio positions and strategies. The performance of a monthly rebalanced optimal portfolio based on these constraints is shown. We also perform a simple test for robustness, noticing that the optimal weights produced in the mean-variance optimal setting are sensitive to initial conditions. This leads us to a variation of the work of Goldfarb and Iyengar in Section 4. There we also provide a formulation similar to their work using an ellipsoidal uncertainty set for the vector of mean returns.

We develop a linear factor model motivated in part by the work of Fama and French. We conduct cross-sectional regressions to fit the regression parameters and note a term structure for these factor returns. This deviates from the regression procedure in Goldfarb and Iyengar. We therefore develop our own variation of uncertainty in input parameters in Section 4.4. For our asset universe, we consider stocks determined dynamically by a lower bound of $10 billion for market capitalization. We avoid survivorship bias as much as possible by selecting all stocks that have traded at least once over the history we are interested in.

The performance of the linear model ex-post is evaluated using a lagged lter. We observe summary portfolio statistics superior to those obtained in a standard mean-variance optimized portfolio constructed in 5. There we also show the results of the robust counterpart. These results provide compelling evidence that, in contrast to the nominal problem, the robust problem in fact delivers a portfolio with intended characteristics; e.g., higher Sharpe with more attractive loss profiles.
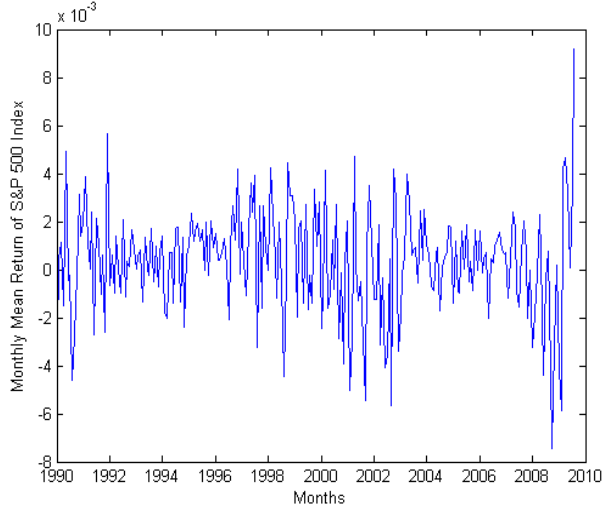
Figure 1: Trailing 200 day mean return of S&P 500 Index.

2 Maximum Sharpe Ratio

The traditional Sharpe ratio optimization problem may be formulated as:

$$\max_{w \in \mathbb{R}^n} \quad \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}} \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^n w_i = 1, \tag{2}$$

$$w \geq 0. \tag{3}$$

Here, we are optimizing over portfolio weights, w, with w_i the weight corresponding to the ith asset. We denote the returns of a risk free asset by r_f. The parameter Σ denotes the covariance matrix of the assets available for the portfolio, and μ represents the mean return. We note that, due to the nonstationarity of financial data, both of these parameters are subject to poor specification. For example, we may consider μ to be the trailing return of the assets over some time window. However, as Figure (1) indicates, this method of determining μ, while common in practice, is neither stable nor accurate when considering future scenarios. A similar statement may be made about correlation and own-variance.

The constraint (3) rules out short selling; that is, we may only buy stocks. Constraint (2) implies that the objective function of the above problem is a homogeneous function of the portfolio w. Assuming that there exists at least one w such that $\mu^T w > r_f$, the normalization condition (2) may be dropped and the constraint $\mu^T w - r_f = 1$ added to the problem to obtain an optimal solution. Using this transformation, the problem reduces to minimizing the worst case variance:

$$\min_{w \in \mathbb{R}^n} \quad w^T \Sigma w \tag{4}$$

$$\text{s.t} \quad \mu^T w - r_f = 1, \tag{5}$$

$$w \geq 0.$$

Notice that the obtained optimal solution of Problem (4) should then be normalized to satisfy the condition (2) and to be an optimal solution of Problem (1). It is worth mentioning that when short selling is not

2

allowed, and $r_i < 0$ for all $i = 1, 2, ..., n$, an optimal solution of Problem (4) is not necessarily an optimal solution of Problem (1), and the transformation does not remain valid. Indeed, negative Sharpe ratio is considered ambiguous and is not typically considered (for a through discussion the reader is referred to [7]). To alleviate negative Sharpe ratios, one may permit short selling when maximizing the Sharpe ratio. Alternatively, one may use other objectives, such as minimizing variance of the return, or heuristic algorithms [6].

Assuming that $r_f \geq -1$, the constraint (5) may be relaxed by $\mu^T w - r_f \geq 1$, since the relaxed constraint will always be tight at an optimal solution. Therefore, we are left to solve the following problem:

$$\min_{w \in \mathbb{R}^n} \quad w^T \Sigma w \tag{6}$$

$$\text{s.t} \quad \mu^T w - r_f \geq 1, \tag{7}$$
$$w \geq 0.$$

When the covariance matrix Σ is positive definite, Problem (6) is a (strictly) convex quadratic programming problem and has a unique solution.

We also consider additional problems by varying the constraints found in Problem (1). For example, investors typically prefer not to invest all of their money in one stock. This is a result of both theory and practice: we see that portfolio risk may be mitigated if true diversification is possible (e.g., in the case where we find assets that retain low correlation even in so-called fear environments), and performance of, say, an even weighted portfolio may often times have more desirable characteristics than an otherwise unconstrained MVO optimal portfolio [1]. Therefore, to obtain a more diversified portfolio we also consider constraining maximum position sizes below, e.g., Problems (9) and (10). In the next section, we describe the portfolios we consider in this preliminary analysis.

In light of the potential for poor specification of the input parameters noted above, we also examine the robustness of the algorithm to perturbations in μ.

3 Analysis of Traditional Sharpe Optimization

To obtain insight about the effectiveness of various portfolios constructed using the methodology described above, we use a universe of five assets made up of global indexes. For this fixed set of assets, we consider four portfolios, described below.

We use daily close prices for five stock indexes from January 1990 to August 2009 to estimate μ and Σ as above. These indexes include the S&P 500 in the United States (SPX), the FTSE 100 (UKX) in London, the French CAC 40, the German DAX, and the Hang Seng (HSI) in Hong Kong. Since each country has different holidays and non-trading days, we face an issue of incomplete return data; if a particular index's record is incomplete in a given day, we could either interpolate this return or simply ignore it. Interpolating the missing return data maintains the sample size; however, it creates bias for our estimation, and could change the correlation between the indexes. Ignoring or deleting an incomplete record, on the other hand, reduces the sample size. In our case, only one to two days out of about twenty trading days were deleted in such months. Taking this into consideration, we choose to eliminate days for which not all indexes have data. We nevertheless ran the tests both, and the results for portfolios constructed where incomplete data was omitted outperformed those portfolios constructed using the interpolation techniques we employed.

Once μ and Σ are determined, we determine optimal weights w_i monthly throughout our historical period by solving the maximum Sharpe ratio problem under various constraints:

Portfolio (P1): Short selling is not allowed and $r_f = 0$:

$$\max_w \frac{\mu^T w}{\sqrt{w^T \Sigma w}} \quad \text{s.t} \quad \sum_{i=1}^n w_i = 1, \quad w \geq 0. \tag{8}$$

Portfolio (P2): Maximum investment on each asset is bounded and a cash account earning LIBOR is introduced so that $r_f \neq 0$:

$$\max_w \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}} \quad \text{s.t} \quad \sum_{i=1}^n w_i = 1, \quad 0 \leq w \leq 0.5. \tag{9}$$

Portfolio (P3): Short selling is allowed and $r_f = 0$:

$$\max_w \frac{\mu^T w}{\sqrt{w^T \Sigma w}} \quad \text{s.t} \quad \sum_{i=1}^n (w_i)_+ = 1, \quad \sum_{i=1}^n (w_i)_- = -1, \quad |w_i| \leq 0.5, \quad i = 1, 2, ..., n. \tag{10}$$

To obtain the optimal solution of Problem (10), we solve the following problem:

$$\min \quad v^T \begin{pmatrix} \Sigma & -\Sigma \\ -\Sigma & \Sigma \end{pmatrix} v, \quad \text{s.t} \quad (\mu, \mu)^T v \geq 1, \quad |v_i| \leq 0.5, \quad \sum_{i=1}^5 v_i = 1, \quad \sum_{i=6}^{10} v_i = -1,$$

or equivalently

$$\min \quad v^T \begin{pmatrix} \Sigma & -\Sigma \\ -\Sigma & \Sigma \end{pmatrix} v, \quad \text{s.t} \quad (\mu, \mu)^T v \geq 1, \quad v_i \leq 0.5 \sum_{i=1}^5 v_i, \quad i = 1, ..., 5, \quad v_i \leq 0.5 \sum_{i=6}^{10} v_i, \quad i = 6, ..., 10. \tag{11}$$

Then the optimal solution of Problem (10) is

$$w^* = \frac{1}{\sum_{i=1}^5 v_i^*} \begin{pmatrix} v_1^* \\ \vdots \\ v_5^* \end{pmatrix} + \frac{1}{\sum_{i=6}^{10} v_i^*} \begin{pmatrix} v_6^* \\ \vdots \\ v_{10}^* \end{pmatrix},$$

where v^* is the optimal solution of Problem (11). We used *CVX: Matlab Software for Disciplined Convex Programming* to solve this problem.

An optimal solution of Problem (10) may be found by solving the following solution:

$$\min_w \quad w^T \Sigma w, \tag{12}$$
$$\mu^T w \geq 1$$
$$\sum_{i=1}^n w_i = 0,$$
$$(w_i)_+ \leq 0.5 \sum_{i=1}^n (w_i)_+, \quad i = 1, 2, ..., n,$$
$$(w_i)_- \leq -0.5 \sum_{i=1}^n (w_i)_-, \quad i = 1, 2, ..., n,$$

$$\tag{13}$$

This is because every optimal solution of Problem (10) is an optimal solution of Problem (12). Conversely, assume that w^* is an optimal solution of Problem (12). Now $w_i = \frac{w_i^*}{\sum (w_i^*)_+}$ if $w_i^* \geq 0$ and $w_i = \frac{w_i^*}{|\sum (w_i^*)_-|}$ if $w_i^* \leq 0$ is an optimal solution of Problem (10). Here we included the constraint $\sum_{i=1}^n w_i = 0$ to guarantee that the total weights for shorts equal the total weights for long trades.

Proposition 3.1. *The optimal solution of the above problem can be obtained by solving the following problem:*

$$\min_w \quad w^T \Sigma w, \tag{14}$$
$$\mu^T w \geq 1$$
$$\sum_{i=1}^{n} w_i = 0,$$
$$w_i \leq t_i \leq 0.5 \sum_{i=1}^{n} t_i, \quad i = 1, 2, ..., n,$$
$$-0.5 \sum_{i=1}^{n} s_i \leq s_i \leq w_i, \quad i = 1, 2, ..., n,$$
$$s_i + t_i = w_i, \quad i = 1, 2, ..., n, \tag{15}$$
$$s_i^2 + t_i^2 = w_i^2, \quad i = 1, 2, ..., n, \tag{16}$$
$$s_i \leq 0, \ t_i \geq 0, \quad i = 1, 2, ..., n.$$

Proof. Let w_i^* be a solution of Problem (12). Then $t_i^* = \max\{0, w_i\}$ and $s_i^* = \min\{0, w_i\}$ satisfies the problem. So (w_i^*, t_i^*, s_i^*) is an optimal solution of Problem (14). Now assume (w_i^*, t_i^*, s_i^*) is an optimal solution of Problem (14). Then constraints (15) and (16) implies that $s_i^* t_i^* = 0$. Since $s_i \leq 0$, $t_i \geq 0$ and $s_i + t_i = w_i$ we must have $t_i^* = \max\{0, w_i\}$ and $s_i^* = \min\{0, w_i\}$. $\qquad\square$

3.1 Performance with Select Index Data

The comparison of the performance of the above strategies is shown in Table (1). These data are obtained using a monthly rebalancing strategy. That is, optimal weights w are computed at the end of every month, and these weights are used to determine performance over the next calendar month. In the first five columns, the performance of the constituent indexes are shown. The last three columns correspond to the investment based on strategies (P1)-(P3) as previously described.

From the table, we can see that the the performance of portfolio P1 is oftentimes closely related to that of Hang Seng. This is a result of the lack of a maximum exposure constraint. In many months, most or all of the weight was put into the HSI for this portfolio. In particular, we note that the optimal weights essentially offer a trend tracking strategy. Also, since only long positions are allowed in this case, a dramatic drop is inventible during 2008. This is relieved to some extent in portfolio P2.

Recall that when computing the portfolio P2, we required that no weight constituted more than 50% of the total portfolio exposure. This was introduced to guarantee some measure of diversity. This condition, along with the availability of a risk free asset produces a portfolio where some risk has been mitigated, and market exposure, i.e., β, has been reduced relative to the unconstrained portfolio.

In portfolio P3, we allow short selling. Theoretically, short positions allow for gains in a portfolio even when the market at large may be dropping. An example of this type of protection is witnessed in 2001. However, during other down years, P3 does not generate positive returns. It does afford some protection, though. Notice that while we don't see gains in 2008 for this portfolio, P3 does have the minimum monthly drawdown among all the portfolios considered; most of which occurred during this dramatic year.

Overall, P1 and P2 do not show significant advantage over the indexes. However, both P1 and P2 produce a positive α relative to the S&P, albeit with significant market exposure. The portfolio constructed using the constraints found in P3, yields a relatively low β, and maintains a (smaller) positive α. This

5

		SPX	HSI	UKX	CAC	DAX	P1	P2	P3
Annual Return (%)	2000	-9.85	-11.00	-10.21	-0.54	-7.54	-12.10	-3.04	-2.86
	2001	-12.06	-24.27	-15.75	-21.97	-19.79	-18.33	-12.73	2.58
	2002	-24.26	-18.89	-25.60	-34.59	-43.94	-28.56	-17.18	-0.39
	2003	26.18	35.10	14.61	16.66	37.08	37.28	18.70	3.16
	2004	9.36	13.07	7.82	8.47	7.34	5.29	3.84	-2.01
	2005	2.86	5.03	16.57	23.18	27.07	5.54	8.55	-11.02
	2006	13.62	34.20	10.71	17.53	21.98	17.84	13.63	0.66
	2007	4.24	37.09	4.12	1.54	22.29	37.91	26.73	25.44
	2008	-39.76	-47.99	-32.18	-42.83	-40.37	-49.28	-31.86	-16.72
	2009	12.91	46.09	6.34	8.06	12.62	20.13	10.82	13.11
Annualized Return (%)	Full Period	-4.03	2.03	-4.17	-5.61	-2.74	-2.76	0.20	0.65
Annualized Sharpe Ratio	2000-2004	-0.15	-0.61	-0.41	-0.33	-0.23	-0.23	-0.16	0.06
	2005-2009	-0.19	0.45	0.01	-0.04	0.34	0.14	0.33	0.18
	Max Gain	9.67	17.07	8.65	13.41	21.38	11.98	10.10	6.68
	Max Drawdown	-16.94	-22.47	-13.02	-17.49	-25.42	-16.39	-12.96	-7.28
S&P Relative	β	1.00	1.05	0.80	1.01	1.23	1.02	0.71	(0.13)
	α	0.00	0.64	-0.07	-0.09	0.30	0.17	0.25	0.06

Table 1: Performance statistics for global indexes and Sharpe optimal strategies. P1 corresponds to a long only portfolio with no riskless asset; P2 to a long only portfolio with a cash account and position limits; and P3 is constructed allowing for short sales and constrained position size. All values except β are in percents.

portfolio, along with P2, also has a positive annualized return over the period under investigation. Among the indexes, only the Hang Seng shows a positive return over this time.

3.2 Sensitivity Analysis

In the optimization procedures, the expected returns μ and covariance Σ are used as parameters. A natural question arises here: is the optimal portfolio obtained from problem (P1) sensitive to the choice of parameter μ? Thinking of this input parameter as a random variable, we investigate how the optimal portfolios behave with respect to perturbations in μ, reflecting possible errors in the estimation of the return vector.

To accomplish this test, we fix a time, and calculate μ and Σ from data. We then generate 2000 instances of a return vector r following a multivariate normal distribution with fixed mean, μ, and fixed covariance, Σ; $r \sim N(\mu, \Sigma)$. For each instance of r, we run the optimization procedure with the input pair (r, Σ) to obtain optimal weights.
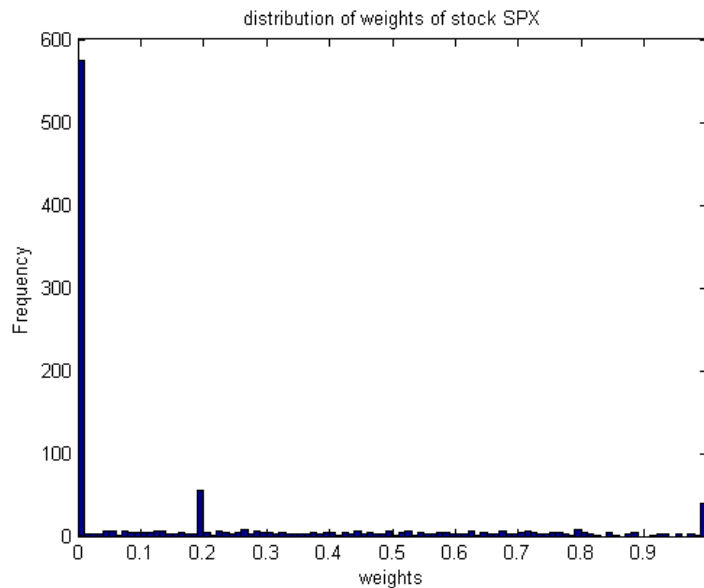

Figure 2: Distribution of optimal weights for the index HSI over input parameters $r \sim N(\mu, \Sigma)$.

We note that, as in Figure (2), the weights for all indexes tend to cluster at zero, one, and the evenly-weighted values. In addition, we notice that the expected weight over all resulting weights from inputs, r differs significantly from the weight obtained using μ and Σ. For example, the average weight for the S&P over all input values r is 0, while the weight using input μ is 0.16. The values for all indexes are found in Table (2)

From this analysis above, we are motivated to use some optimization technique that can take into account uncertainty. We employ a particular robust optimization, produced below, which accounts for both uncertainty in μ as well as the covariance, Σ.

Index	Average of Weights over $r \sim N(\mu, \Sigma)$	Weights Using (μ, Σ)
SPX	0.0000	0.1593
HSI	0.0849	0.1857
UKX	0.6260	0.2979
CAC	0.1309	0.1531
DAX	0.1582	0.2040

Table 2: Comparison of average of optimal weights for the indexes over input parameters $r \sim N(\mu, \Sigma)$ with the weight obtained using μ and Σ.

4   Robust Maximum Sharpe Ratio Problem and a Linear Factor Model

We follow Goldfarb and Iyengar's problem formulation for a robust Sharpe problem [3] to introduce model uncertainty into our optimization procedure. We construct both a linear factor model based in part on the work of Fama and French [2] as well as a new uncertainty set for the vector of expected returns. In addition, we adapt a cross sectional regression methodology to the structure found in [3]. While our adaptation strays from the parsimonious uncertainty sets of Goldfarb and Iyengar, our implementation retains key elements of their method wherever possible.

4.1   Overview of Goldfarb and Iyengar Model and a Variation in Uncertainty Sets

Goldfarb and IyengarHere [3] assume that the return vector follows a linear factor model as in (24):

$$\mathbf{r} = \mu + V^T \mathbf{f} + \epsilon, \tag{17}$$

where μ is the expected return vector, $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, F) \in \mathbb{R}^m$ is the returns of the factors that are assumed to drive market returns, $V \in \mathbb{R}^{m \times n}$ is the factor loading matrix of the n assets and $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{D}) \in \mathbb{R}^n$ is the model residual. Also, we suppose that the residual return vector, ϵ, is independent of the vector of factor returns $\mathbf{f}$, the covariance matrix $F \succ \mathbf{0}$ and the covariance matrix $\mathbf{D} = \mathbf{diag}(d) \succeq \mathbf{0}$. Thus, the vector of asset returns $\mathbf{r} \sim \mathcal{N}(\mu, V^T F V + \mathbf{D})$.

Similar to [3], we assume that the mean return μ, factor loading V, and error covariance D are subject to uncertainty. The uncertainty set S_d for the matrix D is given by

$$S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \ldots, n\} \tag{18}$$

where the individual diagonal elements d_i of $\mathbf{D}$ are assumed to lie in an interval $[\underline{d}_i, \overline{d}_i]$

We assume that the columns of the matrix V belong to the elliptical uncertainty set S_v given by

$$S_v = \{V : V = V_0 + W, \|W_i\|_g \le \rho_i, i = 1, ..., n\}, \tag{19}$$

where W_i is the i-th column of W and $\|w\|_g = \sqrt{w^T G w}$ denotes the elliptic norm of w with respect to a symmetric, positive definite matrix G. In [3], an interval uncertainty set is used to explain the uncertainty in the mean return:

$$\{\mu : \mu = \mu_0 + \xi, |\xi_i| \le \gamma_i, i = 1, ..., n\}, \tag{20}$$

We introduce here the uncertainty set, $\tilde{S}_m$ defined by

$$\tilde{S}_m = \left\{ \mu : \mu = \mu_0 - \sum_{j=1}^{k} u_j \mu^{(j)}, u^T u \leq 1 \right\}. \tag{21}$$

This set is considered so as to provide a method to retain the joint distribution of the returns, r, when allowing uncertainty in our expectation. Specifically, we choose $\mu^{(j)}$ above to be the column vectors of the covariance matrix of returns being modeled, and μ_0 to be the expected return. We note that we do not retain the regression-focused derivation for the uncertainty in μ as in [3].

Using the uncertainty sets S_v, $\tilde{S}_m$ and S_d, the robust maximum Sharpe ratio problem is given by

$$\max_{\{w : w \geq 0, 1^T w = 1\}} \min_{V \in S_v, \mu \in \tilde{S}_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}.$$

We will assume that the optimal value of this max-min problem is strictly positive, i.e., that there is at least one asset with worst case return greater than r_f. The robust Sharpe ratio reduces to minimizing the worst case variance:

$$\min_w \max_{V \in S_v} \quad w^T V^T F V w + w^T \bar{D} w,$$

$$\min_{\mu \in S_m} \quad (\mu - r_f 1)^T w \geq 1. \tag{22}$$

The constraint $\min_{\mu \in S_m} (\mu - r_f 1)^T w$ equals

$$
\begin{aligned}
\min_{\mu \in S_m} (\mu - r_f 1)^T w &= \min_{\mu \in S_m} \left(\mu_0 - \sum_{j=1}^{k} u_j \mu^{(j)} - r_f 1 \right)^T w \\
&= (\mu_0 - r_f 1)^T w - \max_{\{u : \|u\|_2 \leq 1\}} \sum_{j=1}^{k} u_j (\mu^{(j)})^T w \\
&= (\mu_0 - r_f 1)^T w - \sqrt{\sum_{j=1}^{k} ((\mu^{(j)})^T w)^2} = (\mu_0 - r_f 1)^T w - \|Mw\|_2,
\end{aligned}
$$

where the last inequality comes from the definition of dual norms. Thus constraint (25) equals to the following constraint:

$$(\mu_0 - r_f 1)^T w \geq 1 + \|Mw\|_2,$$

where the jth row of the matrix M is $(\mu^{(j)})^T$. Using the discussion in [3], the robust counterpart problem is reduced to the following second order cone programming problem:

$$\min_{\sigma, \tau, t, v, \delta, w} \quad v + \delta \tag{23}$$

$$\text{s.t} \quad \|[2\bar{D}^{\frac{1}{2}} w; 1 - \delta]\| \leq 1 + \delta,$$

$$\|Mw\|_2 \leq (\mu_0 - r_f 1)^T w - 1,$$

$$(\rho^T w; v; w) \in \mathcal{H}(V_0, F, G).$$

where $\mathcal{H}(V_0, F, G)$ is defined as below:

9

Definition 4.1. *Given* $V_0 \in \mathbb{R}^{m \times n}$ *and* $F, G \in \mathbb{R}^{m \times m} \succ 0$, *define* $\mathcal{H}(V_0, F, G)$ *to be the set of all vectors* $(r, v, w) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$ *satisfying the following: there exists* σ, τ *and* $t \in \mathbb{R}^m$ *that satisfy*

$$\sigma, \tau, t \geq 0,$$
$$\tau + 1^T t \leq v,$$
$$\sigma \leq \frac{1}{\lambda_{\max}(H)},$$
$$\|[2r; \sigma - \tau]\| \leq \sigma + \tau,$$
$$\|[2(Q^T H^{\frac{1}{2}} G^{\frac{1}{2}} V_0 w)_i; 1 - \sigma \lambda_i - t_i]\| \leq 1 - \sigma \lambda_i + t_i, \quad i = 1, ..., m$$

where $Q\Lambda Q^T$ *is the spectral decomposition of* $H = G^{-\frac{1}{2}} F G^{-\frac{1}{2}}$, $\Lambda = diag(\lambda_i)$.

Notice that in the derivation of the robust counterpart problem, no assumption is made about nonnegativity of the weights. Further, while the uncertainty sets of Goldfarb and Iyengar are intimately related to the model (24), we note that after the sets S_v, $\tilde{S}_m$ and S_d are determined, the robust problem is not dependent on (24).

Clearly, to construct such robust Sharpe portfolios, we are required to model returns linearly based on factors **f**. Below we construct a linear factor model to implement the program just outlined, with some adjustments.

4.2   Linear Factor Model and Stock Universe

Fama and French in [2] model a cross-section of stock returns based on loadings according to three factors that mimic market exposure, size and value using market beta, market capitalization, and book to price, respectively. Motivated by their work, we employ a similar factor model. In our model, we use asset specific factors:

- Market value: the total value of all common shares currently outstanding.

- Book-to-price: the ratio of the tangible book value to the market value.

- Debt-to-equity: the ratio of the total long term debt to market value.

- Asset growth: the year over year growth in assets as a percent.

- One month trailing return: the return over the most recently completed month.

We calculate factor values monthly based on a universe of stocks described below. We normalize all variables based on the factor specific cross-sectional mean and variance. Additionally, prior to the normalization of market value, we employ a log transformation. The plots in Figure (3) show the scatter plot of monthly returns against market value, before and after a log transformation, respectively. These plot confirms the phenomenon that as the size of a company increases, the variance of returns decreases. We see that taking the log not only greatly reduces the scale of the factor, but also stabilizes the conditional variance. In addition to normalization and transformation, we replace unreported factor values with the factor population mean.

Our universe of stocks is constructed monthly from a Compustat database at the end of each month from August 2000 to May, 2009. To the best of our ability, our implementation does not suffer from survivorship

bias as we include all stocks with any available closing price at any point in our backtest history. Then at each month, we calculate the factors above, excluding those companies that had no trade information in the preceding month. We use quarterly data, and lag all factors but market value by two months to reduce any forward looking bias. We restrict our attention to stocks with a market capitalization over $10 billion, reducing our sample to about 200 stocks a month.
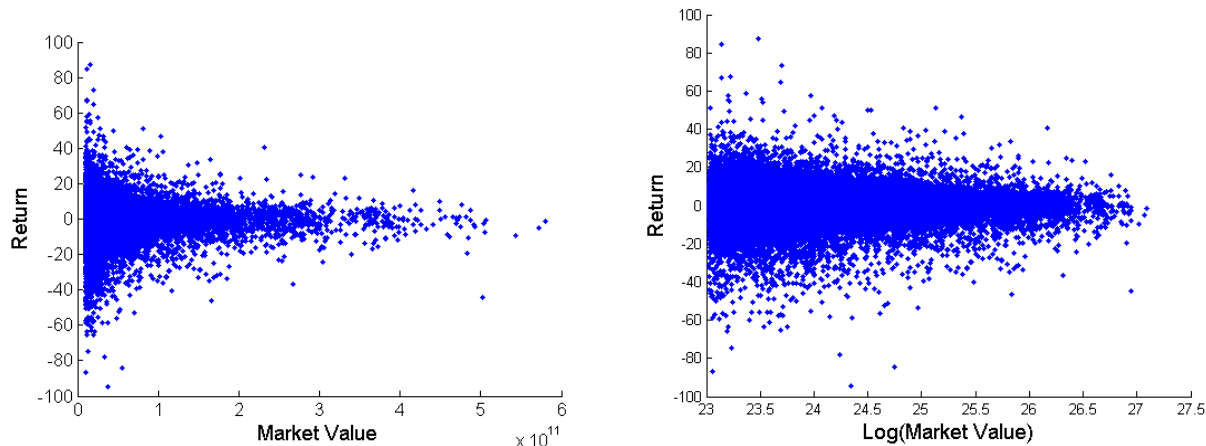

Figure 3: Plots of forward one month returns against market value for our universe of stocks. From left to right, we see the original data and the log-transformed data.

We perform a cross-sectional regression using the five factors above on the one month forward total returns, monthly from August, 2000 to May, 2009. We use an ordinary least square regression with intercept. We have at each time t, then,

$$\mathbf{r}_{t+1} = \mu_t + V_t^T \mathbf{f}_t + \epsilon, \tag{24}$$

where the subscript denotes the time at which the given variable is available as well as emphasizing the time varying nature of all the terms.

4.3 Analysis of Factor Model

To evaluate our factor model, we follow the methodology in Haugen [4] as well as reexamining the portfolio construction method in problem P3 above with a universe of stocks rather than index data. For our comparison to P3, we look at using a trailing historical mean for μ as well as using a model forecasted μ.

We follow a portfolio construction as in [4]. After performing the cross sectional regressions outlined above, we use lagged 12 and 18 month filters on the calculated factor returns, $\mathbf{f}_t$, to construct a forecast for returns. We then construct evenly weighted portfolios of the top and bottom deciles based on these forecasts. We then synthesize being long the top decile portfolio and short the lowest decile portfolio.

Some performance metrics for the 12 and 18 month filter forecasted long-short portfolios are found in Table 3. We note the superior performance of this heuristic portfolio methodology to the mean-variance optimization portfolios built from the indexes (as in Table 1). In almost every year, and in nearly every metric we report, both the 12 and 18 month filter outperform the portfolio, P3, above. We note that this is not an entirely fair comparison as we are comparing an asset based basket to a very small index based basket. We construct the asset based equivalent of P3 below using historical returns. There we will see the

11

		μ: 12 Month Filter	μ: 18 Month Filter
Annual	2002 (beginning March)	20.25	-0.94
Return (%)	2003	3.24	13.13
	2004	24.90	24.51
	2005	17.13	18.43
	2006	2.40	8.71
	2007	35.91	24.65
	2008	-9.62	-9.36
	2009 (through June)	-8.07	-2.24
Annualized Return (%)	Full Period	10.65	9.79
Annualized Volatility (%)	Full Period	16.28	15.44
Annualized Sharpe Ratio		0.70	0.69
	Max Gain	13.39	10.63
	Max Drawdown	-14.32	-14.62
Relative to S&P	β	(0.02)	(0.08)
	α	0.95	0.87

Table 3: Performance metrics for 12 and 18 month filtering.

18 month filter outperforming in some (most) years, and in some metrics, but the results are in no way as conclusive as the comparison to the index based optimization. We choose to focus on the 18 month filter in what follows, but note that the same analysis may be performed with the 12 month version.

In Figure 4, we show the factor returns as well as the 18 month filter. We notice that each factor return has a term structure and that throughout 2008 and into 2009 there is a marked change in some variables, both in magnitude and sign. In particular, we note the abrupt change in our estimate of the intercept, signaling a reversal from the prior bull market.

4.4   Uncertainty Sets Used in Implementation

Goldfarb and Iyengar in [3] construct uncertainty sets based on multivariate linear regression methods and results. There they assume that each asset has its own factor return vector $\mathbf{f}$, using a trailing history to obtain multiple observations. We employ a cross sectional regression; that is, we take assets at a specific time as our multiple observations. This materially impacts the uncertainty set construction they suggest. While we do not maintain their construction methodology, we adapt their elliptical uncertainty sets to our setting. This procedure of using uncertainty sets that a practitioner may deem useful (at least heuristically) is more in the tradition of [8].

We consider the problem

$$\min_{w} \quad \max_{V \in S_v} \quad w^T V^T F V w + w^T \bar{D} w,$$
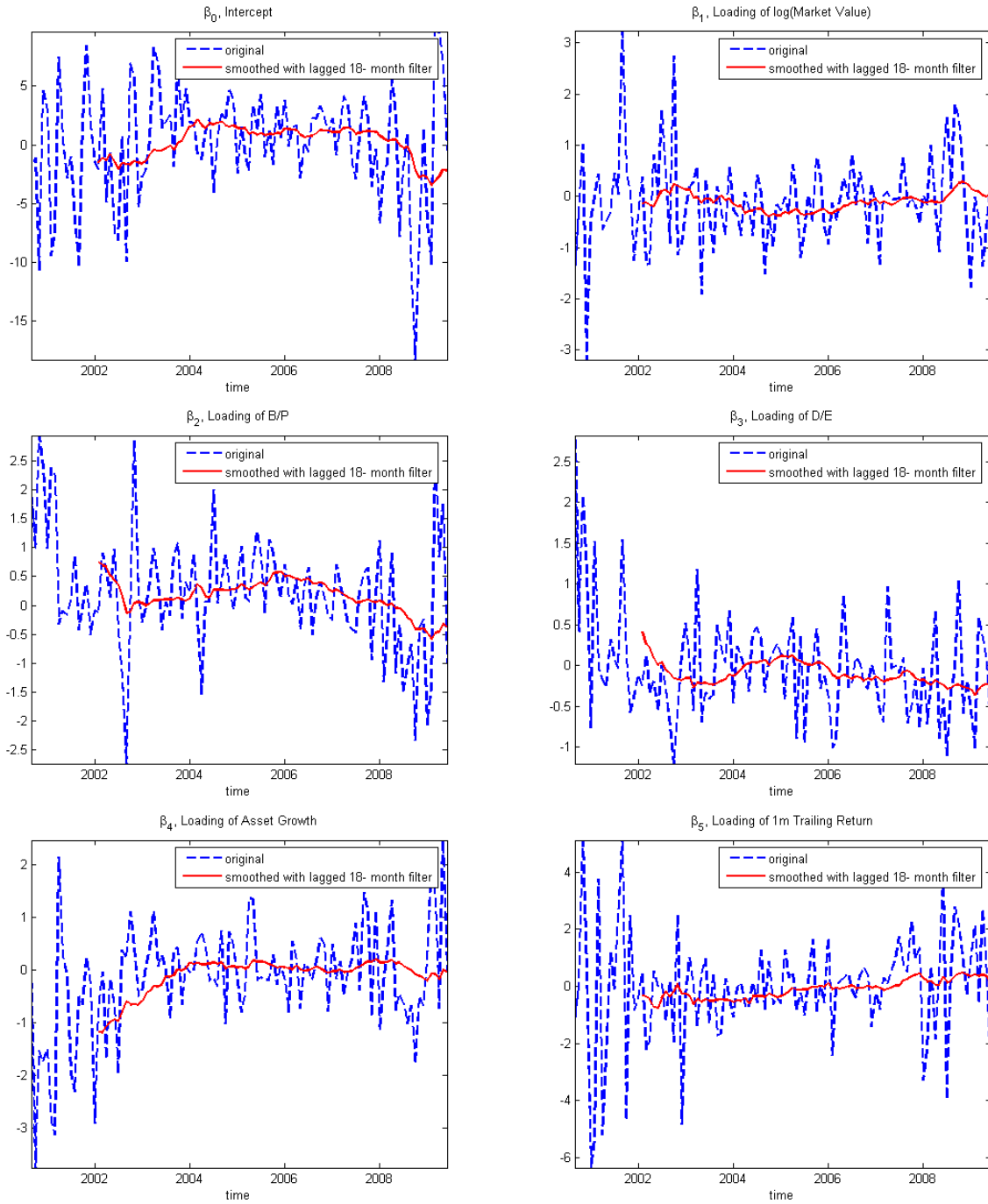$$\min_{\mu \in S_m} \quad (\mu - r_f 1)^T w \geq 1, \tag{25}$$

Figure 4: From top left to bottom right, we see the term structure of the factor returns for μ_0 (intercept), market value, book-to-price, debt-to-equity, asset growth, and trailing returns. In each figure, the solid line is a plot of the 18 month filter.

where

$$S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \dots, n\},$$

$$S_v = \{V : V = V_0 + W, \|W_i\|_g \leq \rho_i, i = 1, \dots, n\},$$

$$S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leq \gamma_i, i = 1, \dots, n\}.$$

Recall that this was the problem derived from

$$\max_{w \in \mathcal{W}} \min_{V \in S_v, \mu \in S_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}},$$

where the return vector is assumed to follow (24), and $\mathcal{W}$ is some constraint set on the weights w. Further, this problem may be written as a second order cone program as in (23), with slight modification.

We are left then to determine the bounds $\overline{d}_i$, ρ_i, and γ_i, for $i = 1, \dots, n$, and to define the elliptical norm in the set definition for S_v. While our sets are not entirely motivated by regression uncertainty, we make only slight modifications. The interested reader may see [3] for the initial development of what follows.

We begin by choosing to use 18 months of trailing factor returns, $\mathbf{f}_t$, for what follows. Since our modeling is cross-sectional, we look temporally for the variance of factor returns, F. We define the matrix B as

$$B = [\mathbf{f}_{t-p} \ \cdots \ \mathbf{f}_{t-2} \ \mathbf{f}_{t-1}]. \tag{26}$$

with $p = 18$. For ease of notation, we include the intercept term from above in $\mathbf{f}_{cdot}$ here and for the remainder. This gives the maximum likelihood estimate for the covariance of factor returns:

$$F = \frac{1}{p-1} \cdot \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T \right), \tag{27}$$

with $\mathbf{1}$ a vector of ones with p rows. We define G as $G = (p-1) \cdot F$, and the elliptical norm $\|\cdot\|_g$ is given by $\|x\|_g = \sqrt{x^T G x}$. We choose

$$\overline{d}_i \equiv d, \tag{28}$$

where d is the worst case variance of errors in the return over the trailing p periods. The values for ρ_i and γ_i are chosen to be based on a confidence level ω, under an F-distribution with $p - m - 1$ degrees of freedom, with m the number of factors in the model. Specifically, let $c_{p,m}(\omega)$ denote such a critical value. Then we let

$$\gamma_i = \sqrt{(B \cdot B^T)_{(1,1)}^{-1} c_{p,m}(\omega) \cdot \sigma_i^2}, \tag{29}$$

where σ_i^2 is the own-variance of asset i over the trailing history of p periods, scaled to the correct time units for forecasting. We use

$$\rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}. \tag{30}$$

At each month over our backtest history, we calculate the above parameters, and solve the robust Sharpe problem using $\omega = 95\%$. At time t, the input design matrix is simply V_t, and we use the forecasted μ_0 given by

$$\begin{aligned} \mu_0 &= E([\mathbf{1} \ V_t]^T \mathbf{f}_t) \\ &= [\mathbf{1} \ V_t]^T E(\mathbf{f}_t). \end{aligned} \tag{31}$$

We approximate $E(\mathbf{f_t})$ by $\frac{1}{p} \sum_{l=1}^p (\mathbf{f}_{t-l})$.

5 Analysis of Robust Portfolio Optimization: Performance and Comparison

We present performance results for solving the above robust Sharpe problem under the constraints set out in P3 above. Namely, maximum exposure (reduced here to five percent), and short sales allowed with a dollar neutral portfolio. We note that we initially considered problem P2, but approximately thirty percent of months in our sample were such that the constraints based on γ_i gave a null feasible set. That is, the worst case expected performance oftentimes was below the risk-free rate for all stocks under consideration. We do not see this as a drawback of the procedure, however, as this could be relieved by considering a variance minimization rather than Sharpe maximization, say. We retain our focus on only the Sharpe problem, though, and therefore only consider the robust asset based counterpart of P3.

We rebalance a theoretical portfolio monthly, calculating the performance of the robust portfolio with weights obtained from solving our problem based on the linear factor model (24). We use the software, Mosek with a Matlab interface to solve the problem. We also consider the direct analogy of P3 using the assets available at each month (rather than index data as in our original formulation), both by replacing μ with its model forecasted value as well as using the trailing mean return.

When considering the nominal problem, we use two years of trailing return data to construct the covariance matrix. However, our universe has approximately $N = 200$ assets, so that the number of observations is far fewer than the number of unknown parameters in the covariance matrix. A naive covariance calculation in such a setting may yield ill-conditioned covariance matrices. We remedy this by using the methodology outlined in [5]. There they provide a shrinkage estimator, calculating an optimally weighted average of the sample covariance matrix and the identity matrix. The method provides a well-conditioned, invertible covariance matrix that is also a consistent estimator.

The results of our analysis our found in Table 4. We notice that on an annualized basis each of the actively constructed portfolios outperforms the S&P 500 Index, with little β exposure and positive, statistically significant monthly α. We have, however, neglected transaction costs, but these should be to a minimum given the universe under consideration.

The more volatile, traditional Sharpe optimized portfolio using a trailing mean for μ loses a considerable amount through the middle of 2009. This is likely due to the dramatic reversal in the markets after March of that year, and the tendency that we noted earlier for such a portfolio to be trend following. We see considerable outperformance in terms of annualized return (7.37% annualized return) from the nominal problem using a forecasted return. Further, we see a reduction of annualized volatility relative to the traditional trailing mean based portfolio. However, the volatility of the forecasted μ, Sharpe optimized portfolio is approximately the same as the market index.

The robust Sharpe portfolio using the 18 month filter forecast for μ has a low annualized volatility of just 6.19% from 2002-2009. This is coupled with a lower annualized return (relative to the nominal with forecast) of 4.31%. Additionally, this portfolio provides the highest Sharpe ratio among those portfolios considered here with a value of 0.71. However, this is essentially the same result obtained for the naive portfolio created using this same μ (viz., Table 3). This heuristic construction equals or outperforms the traditional Sharpe optimized portfolio, and provides the most α of all strategies considered. It has a significantly negative year in 2008, however.

It is interesting to note that every one of the optimal portfolios considered here is positive in 2008, a disastrous year for many equity based investments. Further, the robust portfolio does not see a loss exceeding 5% in any month in our history. This is balanced by not having any outsized gains, either. We note that this is the intended outcome for such a portfolio.

		S&P 500	Nominal Sharpe μ: Trailing Mean	Nominal Sharpe μ: 18 Month Filter	Robust Sharpe μ: 18 Month Filter
Annual Return (%)	2002 (beginning March)	-17.15	0.33	-0.34	0.21
	2003	13.02	5.20	4.92	2.10
	2004	10.93	1.03	2.64	11.01
	2005	6.45	24.94	5.62	7.30
	2006	12.10	-10.69	-0.27	4.55
	2007	5.75	57.13	21.78	-0.49
	2008	-39.49	2.49	3.36	8.78
	2009 (through June)	2.56	-26.52	18.38	-1.21
Annualized Return (%)	Full Period	-2.78	4.78	7.37	4.31
Annualized Volatility	Full Period	15.78	24.28	15.93	6.19
Annualized Sharpe Ratio		(0.10)	0.32	0.53	0.71
	Max Gain	9.39	18.86	15.85	6.82
	Max Drawdown	-16.94	-26.69	-14.36	-4.85
S&P Relative	β	1.00	(0.11)	(0.09)	(0.06)
	α	0.00	0.63	0.69	0.36

Table 4: Performance statistics for dynamic universe of stocks based on lower market cap bound of \$10 billion. The performance of the S&P is provided as a reference. Each portfolio is constructed to be dollar neutral, with position limits of five percent. The nominal problem is the standard mean-variance optimization problem. All values except β are in percents.

Finally, we note that in the end, likely due in part to nonstationarity of covariance, the traditional Sharpe optimized portfolios (using both a forecast and trailing mean) produce results with lower Sharpe ratios than a simple heuristic method. We believe this is remedied in the robust case, where the methodology contributes in a meaningful way to realistic portfolio construction.

6 Concluding Remarks and Future Work

We tested the performance of the traditional Sharpe optimization problem, and noted both the nonstationarity of the underlying input parameters as well as the sensitivity of the procedure to perturbations in those same inputs. We addressed the latter of these issues using the work of Goldfarb and Iyengar to formulate a robust counterpart to the original problem.

We developed a linear factor model, and conducted cross-sectional regressions to fit the model parameters. These regression outputs were found to have a term structure. We examined a heuristic portfolio construction procedure using lagged filters on the model factor returns to produce forecasted returns. These portfolios outperformed the index-based optimized portfolios, and we decided on an 18 month filter based on its performance. This model was then used to construct uncertainty sets for a robust Sharpe problem. Additionally, we provided another formulation of a robust Sharpe problem, suggesting a new uncertainty set for the expected return. We did not solve this problem, though.

Our method deviated from Goldfarb and Iyengar in that our regressions were cross-sectional and not temporal. Therefore, their derivations of their uncertainty sets did not obtain in our case. We produced modifications to adapt to our regressions. However, we did not rigorously derive or develop said sets.

The performance of the robust Sharpe portfolio using our five-factor linear model was shown to produce very desirable results: very minimal losses, with an attractive Sharpe and positive α. All of the long-short optimal portfolios we considered generated positive returns over the period 2002-2009. The traditional Sharpe problem using a trailing mean was the least attractive of the portfolios we considered in almost every summary statistic we reported: annualized return, volatility, maximum gain, maximum drawdown, Sharpe ratio, and α.

In future work, we would like to further our development of the uncertainty sets used above. That is, we would like to produce regression-based uncertainty sets when considering a cross-sectional setting. Further, we would like to implement the new robust problem we formulated. This new problem is attractive in that it seeks to maintain the correlation structure between returns when considering return uncertainty.

References

[1] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.

[2] E. Fama and K. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1992.

[3] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

[4] R. A. Haugen and N. L. Bakerb. Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41:401–439, 1996.

[5] O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88:365–411, 2004.

[6] D. N. Nawrocki. Portfolio optimization, heuristics, and the "butterfly effect". *Journal of Financial Planning*, pages 68–78, 2000.

[7] J. D. Schwager. *Managed Trading: Myths and Truths*. Wiley; 1 edition, 1996.

[8] M. Koenig and R. H. Ttnc. Robust asset allocation. *Annals of Operations Research*, 132:157–187, 2004.