Parametric Portfolio Policies:
Exploiting Characteristics in the Cross Section of
Equity Returns[*]

Michael W. Brandt
Fuqua School of Business
Duke University[†]
and NBER

Pedro Santa-Clara
Anderson School
UCLA[‡]
and NBER

Rossen Valkanov
Anderson School
UCLA[§]

March 2005

Abstract

We propose a novel approach to optimizing portfolios with large numbers of
assets. We model directly the portfolio weight in each asset as a function of the
asset's characteristics. The coefficients of this function are found by optimizing
the investor's average utility of the portfolio's return over the sample period.
Our approach is computationally simple, easily modified and extended, produces
sensible portfolio weights, and offers robust performance in and out of sample.
In contrast, the traditional approach of first modeling the joint distribution of
returns and then solving for the corresponding optimal portfolio weights is not
only difficult to implement for a large number of assets but also yields notoriously
noisy and unstable results. Our approach also provides a new test of the portfolio
choice implications of equilibrium asset pricing models. We present an empirical
implementation for the universe of all stocks in the CRSP-Compustat dataset,
exploiting the size, value, and momentum anomalies.

1   Introduction

Stock characteristics, such as the firm's market capitalization, book-to-market ratio, or lagged return, are related to the stock's expected return, variance, and covariance with other stocks.[1] However, exploiting this fact in portfolio management has been, up to now, extremely difficult. The traditional mean-variance approach of Markowitz (1952) requires modeling the expected returns, variances, and covariances of all stocks as functions of their characteristics. This is not only a formidable econometric problem given the large number of moments involved and the need to ensure the positive definiteness of the covariance matrix, but the results of the procedure are also notoriously noisy and unstable (e.g., Michaud, 1989). In practice, the Markowitz approach is therefore implemented along with a number of different fixes, including shrinkage of the estimates, imposing a factor structure on the covariance matrix, estimation of expected returns from an asset pricing model, or constraining the portfolio weights.[2] While these fixes generally improve the properties of the optimized portfolio, they require substantial resources such as the tools developed by BARRA, Northfield, and other companies. As a result, formal portfolio optimization based on firm characteristics is seldom implemented by asset managers (with the notable exception of quant managers which are a small part of the profession), even though it has the potential to provide large benefits to the investors.[3]

We propose a simple new approach to equity portfolio optimization based on firm characteristics. We parameterize the portfolio weight of each stock as a function of the firm's characteristics and estimate the coefficients of the portfolio policy by maximizing the utility that would have been obtained by implementing the policy over the sample period.

Our approach has a number of conceptual advantages. First, we avoid completely the auxiliary, yet very difficult, step of modeling the joint distribution of returns and characteristics and instead focus directly on the object of interest — the portfolio weights. Second, parameterizing the portfolio policy leads to a tremendous reduction in dimensionality. For a problem with N stocks, the traditional Markowitz approach requires

[1]Fama and French (1996) find that these three characteristics robustly describe the cross-section of expected returns. Chan, Karceski, and Lakonishok (1998) show that these characteristics are also related to the variances and covariances of returns out-of-sample.

[2]See Black and Litterman (1992), Chan, Karceski, and Lakonishok (1999), Frost and Savarino (1986, 1988), Jagannathan and Ma (2002), Jobson and Korkie (1980, 1981), Jorion (1986), Ledoit and Wolf (2003a, 2003b), Pastor (2000), and Pastor and Stambaugh (2000, 2002). Brandt (2004) surveys the literature.

[3]See for instance Chan, Karceski, and Lakonishok (1999) and Jagannathan and Ma (2002).

modeling N first and $(N^2 + N)/2$ second moments of returns. With preferences other than the simplistic quadratic utility, the traditional approach involves a practically unmanageable number of higher moments for even a relatively small number of stocks (e.g., 100 stocks have over 300,000 third moments). In contrast, our approach involves modeling only N portfolio weights regardless of the investor's preferences and the joint distribution of asset returns. Because of this reduction in dimensionality, our approach escapes the common statistical problems of imprecise coefficient estimates and overfitting, while allowing us to solve very large-scale problems with arbitrary preferences. Third, but related, our approach captures implicitly the relation between the characteristics and expected returns, variances, covariances, and even higher order moments of returns, since they affect the distribution of the optimized portfolio's returns and therefore the investor's expected utility. Fourth, by framing the portfolio optimization as a statistical estimation problem with an expected utility objective function (a "maximum expected utility" estimator as opposed to the usual least-squares or maximum likelihood estimators), we can easily test individual and joint hypotheses about the optimal portfolio weights.

From a practical perspective, our approach is simple to implement and produces robust results in and out of sample. It is also easily modified and extended. We discuss a number of possible extensions, including the use of different objective functions, the use of different parameterizations of the portfolio policy to accommodate short-sale constraints, accounting for interactions between the firm characteristics, conditioning the portfolio policy on macroeconomic predictors, and multi-period investment horizons.

Our paper is related to a recent literature on drawing inferences about optimal portfolio weights without explicitly modeling the underlying return distribution. Brandt (1999) and Ait-Sahalia and Brandt (2002) model the optimal allocations to stocks, bonds, and cash as nonparametric functions of variables that predict returns. Nigmatullin (2003) extends their nonparametric approach to incorporate parameter and model uncertainty in a Bayesian setting. More closely related to our paper is Brandt and Santa-Clara (2004), who study a market-timing problem involving stocks, bonds, and cash by modeling the optimal portfolio weights as functions of the predictors. Specifically, they model the weight in each asset class as a separate function (with coefficients that are specific to the asset class) of a common set of macroeconomic variables. Their approach is relevant for problems involving a few assets that have fundamentally different characteristics, such as the allocation of capital across different asset classes. In contrast, our paper models the weight invested in each

asset as the same function (with common coefficients) of asset-specific variables. This is the relevant problem when choosing among a large number of essentially similar assets, such as the universe of stocks.

Our approach also has a positive use for testing the portfolio implications of equilibrium in asset markets. More concretely, we specify the optimal portfolio weight in each stock as the sum of its market capitalization weight and an optimal deviation from that market cap weight that depends parametrically on the characteristics of the firms. A test of whether the optimal deviations from the market cap weights are jointly zero therefore addresses the hypothesis of whether a representative investor with the given preferences optimally holds the market portfolio. Failure to reject this hypothesis identifies the investor's preferences as consistent with equilibrium in asset markets.

We use our approach to optimize a portfolio of all the stocks in the CRSP/Compustat dataset from 1964 through 2002, using as characteristics the market capitalization, book-to-market ratio, and lagged one-year return of each firm. The investor is assumed to have constant relative risk aversion (CRRA) preferences. Our empirical results document the importance of the firm characteristics for explaining deviations of the optimal portfolio weights from observed market capitalization weights. Relative to market cap weights, the optimal portfolio with and without short-sale constraints allocates considerably more wealth to stocks of small firms, firms with high book-to-market ratios (value firms), and firms with high lagged returns (winners). With a relative risk-aversion of five, the certainty equivalent gain from investing in the optimal portfolio relative to holding the market is an annualized 10% without short-sale constraints and 3% with short-sale constraints. In the case without short-sale constraints, the benefits are even greater when we take into account interactions between the characteristics, especially interactions between lagged returns and book-to-market ratios, and when we allow the coefficients of the portfolio policy to depend on the slope of the yield curve (certainty equivalent gains of 18% and 12%, respectively). In contrast, with short-sale constraints, these two extensions yield smaller economic benefits relative to the base case. We use an out-of-sample experiment to show that the gains in expected utility are not the result of in-sample overfitting.

The remainder of the paper proceeds as follows. We describe the basic idea and various extensions of our approach in Section 2. The empirical application is presented in Section 3. We conclude in Section 4.

2 Methodology

2.1 Basic Idea

Suppose that at each date t there is a large number, N_t, of stocks in the investable universe.[4] Each stock i has a return of $r_{i,t+1}$ from date t to $t+1$ and is associated with a vector of firm characteristics $x_{i,t}$ observed at date t. For example, the characteristics could be the market capitalization of the stock, the book-to-market ratio of the stock, and the lagged twelve-month return on the stock. The investor's problem is to choose the portfolio weights $w_{i,t}$ to maximize the conditional expected utility of the portfolio's return $r_{p,t+1}$:

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} \mathrm{E}_t\big[u(r_{p,t+1})\big] = \mathrm{E}_t\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right]. \tag{1}$$

We parameterize the optimal portfolio weights as a function of the stocks' characteristics:

$$w_{i,t} = f(x_{i,t}; \theta). \tag{2}$$

In most of the paper, we concentrate on the following simple linear specification for the portfolio weight function:

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t} \theta^\top \hat{x}_{i,t} \tag{3}$$

where $\bar{w}_{i,t}$ is the weight of stock i at date t in a benchmark portfolio such as the value-weighted market portfolio, θ is a vector of coefficients to be estimated, and $\hat{x}_{i,t}$ are the characteristics of stock i, standardized cross-sectionally to have zero mean and unit standard deviation across all stocks at date t. Note that, rather than estimating one weight for each stock, we estimate weights as a single function of characteristics that applies to all stocks.

This particular parameterization captures the idea of active portfolio management relative to a performance benchmark. The intercept is the weight of the stock in the benchmark portfolio and the term $\theta^\top \hat{x}_{i,t}$ represents the deviations of the optimal portfolio weight from this benchmark. The characteristics are standardized for two reasons. First, the cross-sectional distribution of the standardized $\hat{x}_{i,t}$ is stationary through time, while that of the raw $x_{i,t}$ may be non-stationary. Second, the standardization implies that the cross-

---

[4]Our method automatically accommodates the realistic case of a varying number of stocks through time. This is not trivially done in the traditional approach as discussed by Stambaugh (1997).

sectional average of $\theta^\top \hat{x}_{i,t}$ is zero, which means that the deviations of the optimal portfolio weights from the benchmark weights sum to zero, and therefore that the optimal portfolio weights always sum to one. Finally, the term $1/N_t$ is a normalization that allows the portfolio weight function to be applied to an arbitrary and time-varying number of stocks. Without this normalization, doubling the number of stocks without otherwise changing the cross-sectional distribution of the characteristics results in twice as aggressive allocations, even though the investment opportunities are fundamentally unchanged.

The most important aspect of our parameterization is that the coefficients θ are constant across assets and through time. Constant coefficients across assets implies that the portfolio weight in each stock depends only on the stock's characteristics and not on the stock's historic returns. Two stocks that are close to each other in characteristics associated with expected returns and risk should have similar weights in the portfolio even if their sample returns are very different. The implicit assumption is that the characteristics fully capture all aspects of the joint distribution of returns that are relevant for forming optimal portfolios. Constant coefficients through time means that the coefficients that maximize the investor's conditional expected utility at a given date are the same for all dates and therefore also maximize the investor's unconditional expected utility.

These two facts imply that we can rewrite the conditional optimization with respect to the portfolio weights $w_{i,t}$ in equation (1) as the following unconditional optimization with respect to the coefficients θ:

$$\max_\theta \ \mathrm{E}\big[u(r_{p,t+1})\big] = \mathrm{E}\left[u\left(\sum_{i=1}^{N_t}\left(\bar{w}_{i,t} + \frac{1}{N_t}\theta^\top \hat{x}_{i,t}\right)r_{i,t+1}\right)\right]. \tag{4}$$

We can then estimate the coefficients θ by maximizing the corresponding sample analogue:

$$\begin{aligned}
\max_\theta \ \frac{1}{T}\sum_{t=0}^{T-1} u(r_{p,t+1}) &= \frac{1}{T}\sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right) \\
&= \frac{1}{T}\sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t}\left(\bar{w}_{i,t} + \frac{1}{N_t}\theta^\top \hat{x}_{i,t}\right)r_{i,t+1}\right),
\end{aligned} \tag{5}$$

for some pre-specified utility function (e.g., quadratic or CRRA utility).

Three observations about our approach are worth making at this point. First, optimizing a portfolio of a very large number of stocks is extremely simple. Given the

relatively low dimensionality of the parameter vector, it is computationally trivial to optimize the portfolio with nonlinear optimization methods.[5] The computational burden of our approach only grows with the number of characteristics entering the portfolio policy, not with the number of assets in the portfolio. Second, the formulation is numerically robust. We optimize the entire portfolio by choosing only a few parameters θ. This parsimony reduces the risk of in-sample overfitting since the coefficients will only deviate from zero if the respective characteristics offer an interesting combination of return and risk consistently across stocks and through time. For the same reason, the optimized portfolio weights tend not to take extreme values. Third, the optimization takes into account the relation between the characteristics and expected returns, variances, covariances, and even higher order moments of returns, to the extent that they affect the distribution of the optimized portfolio's returns and therefore the investor's expected utility. In the optimization, the degree of cross-sectional predictability of each component of the joint return distribution is intuitively weighted by its impact on the overall expected utility of the investor.

To better understand this third point, we can approximate the expected utility of the investor with a Taylor series expansion around the portfolio's expected return $\mathrm{E}[r_{p,t+1}]$:

$$
\begin{aligned}
\mathrm{E}\big[u(r_{p,t+1})\big] \approx\ & u(\mathrm{E}[r_{p,t+1}]) + \frac{1}{2}u''(\mathrm{E}[r_{p,t+1}])\mathrm{E}\big[(r_{p,t+1} - \mathrm{E}[r_{p,t+1}])^2\big] \\
& + \frac{1}{6}u'''(\mathrm{E}[r_{p,t+1}])\mathrm{E}\big[(r_{p,t+1} - \mathrm{E}[r_{p,t+1}])^3\big] + ...
\end{aligned}
\tag{6}
$$

This expansion shows that, in general, the investor cares about all the moments of the distribution of the portfolio return. Since the portfolio return is given by:

$$
r_{p,t+1} = \sum_{i=1}^{N_t} \left(\bar{w}_{i,t} + \frac{1}{N_t}\theta^\top \hat{x}_{i,t} \right) r_{i,t+1}
\tag{7}
$$

the moments of its distribution depend implicitly on the joint distribution of the returns and characteristics of all firms. The coefficients θ affect the distribution of the portfolio's return by changing the weights given to the returns of the individual firms in the overall portfolio.

To perform a comparable portfolio optimization using the traditional Markowitz approach requires modeling the means, variances, and covariances of all the stocks as functions of their characteristics. This entails estimating for each date t a large number of N_t

---

[5]Especially because, for utility functions commonly used and given the linearity of the portfolio policy in the coefficients θ, it is easy to derive analytically the gradient and the Hessian of the optimization problem.

conditional expected returns and $(N_t^2 + N_t)/2$ conditional variances and covariances. Besides the fact that the number of these moments grows quickly with the number of stocks, making robust estimation a real problem, it is extremely challenging to estimate the covariance matrix as a function of stock characteristics in a way that guarantees its positive definiteness. Furthermore, extending the traditional approach beyond first and second moments, when the investor's utility function is not quadratic, is practically impossible because it requires modeling not only the conditional skewness and kurtosis of each stock but also the numerous high-order cross-moments.

Finally, when the benchmark is the value-weighted market, m, the return of the portfolio can be written as:

$$r_{p,t+1} = \sum_{i=1}^{N_t} \bar{w}_{i,t} \, r_{i,t+1} + \sum_{i=1}^{N_t} \left(\frac{1}{N_t} \theta^\top \hat{x}_{i,t} \right) r_{i,t+1} = r_{m,t+1} + r_{h,t+1} \tag{8}$$

where h is a long-short hedge fund with weights $\theta^\top \hat{x}_{i,t}/N_t$ that add up to zero. Therefore problem (6) can be reinterpreted as the problem of a hedge fund that optimizes its portfolio to maximize the utility of investors who already hold the market.

2.2   Statistical Inference

By formulating the portfolio problem as a statistical estimation problem, we can easily obtain standard errors for the coefficients of the weight function. The "maximum expected utility" estimator $\hat{\theta}$, defined by the optimization problem (5), satisfies the first-order conditions:

$$\frac{1}{T} \sum_{t=0}^{T-1} h(r_{t+1}, x_t; \theta) \equiv \frac{1}{T} \sum_{t=0}^{T-1} u'(r_{p,t+1})(x_t^\top r_{t+1}) = 0 \tag{9}$$

and can therefore be interpreted as a method of moments estimator. From Hansen (1982), the asymptotic covariance matrix of this estimator is:

$$\Sigma_\theta \equiv \text{AsyVar}[\hat{\theta}] = \frac{1}{T} [G^\top V^{-1} G]^{-1}, \tag{10}$$

where

$$G \equiv \frac{1}{T} \sum_{t=0}^{T-1} \frac{\partial h(r_{t+1}, x_t; \theta)}{\partial \theta} = \frac{1}{T} \sum_{t=0}^{T-1} u''(r_{p,t+1})(x_t^\top r_{t+1})(x_t^\top r_{t+1})^\top \tag{11}$$

7

and V is a consistent estimator of the covariance matrix of $h(r, x; \theta)$.

Assuming marginal utilities are uncorrelated, which is true by construction when the portfolio policy is correctly specified and the optimization is unconstrained, we can consistently estimate V by:

$$\frac{1}{T} \sum_{t=0}^{T-1} h(r_{t+1}, x_t; \hat{\theta}) h(r_{t+1}, x_t; \hat{\theta})^{\top}. \tag{12}$$

If we want to allow for the possibility of a misspecified portfolio policy (e.g., for the purpose of specification testing discussed further below) or if constraints are imposed, we may instead use an autocorrelation-adjusted estimator of V (e.g., Newey and West, 1987).

Alternatively, the covariance matrix of coefficients $\hat{\Sigma}_\theta$ can be estimated by bootstrap. For that, we simply generate a large number of samples of returns and characteristics by randomly drawing monthly observations from the original data set (with replacement).[6] For each of these bootstrapped samples, we estimate the coefficients of the optimal portfolio policy and compute the covariance matrix of the coefficients across all the bootstrapped samples. This approach has the advantage of not relying on asymptotic results and takes into account potentially non-normal features of the data. The bootstrapped standard errors are particularly appropriate in the multi-period investment horizon setup discussed below.

The resulting estimate of the covariance matrix of the coefficients $\hat{\Sigma}_\theta$ can be used to test individual and joint hypotheses about the elements of θ. These tests address the economic question of whether a given characteristic is related to the moments of returns in such a way that the investor finds it optimal to deviate from the benchmark portfolio weights according to the realization of the characteristic for each stock. It is important to recognize that this is not equivalent to testing whether a characteristic is cross-sectionally related to the conditional moments of stock returns for at least two reasons. First, the benchmark portfolio weights may already reflect an exposure to the characteristics and it may not be optimal to change that exposure. Second, a given characteristic may be correlated with first and second moments in an offsetting way, such that the conditionally optimal portfolio weights are independent of the characteristic.

The interpretation of our approach as a method of moments estimator suggests a way

---

[6]We also experimented with block bootstrapping techniques that maintain the time-series dependence of the data (e.g., Politis and Romano (1994)). The resulting inferences are qualitatively the same.

of testing the functional specification of the portfolio policy. In going from equation (1) to equation (4) we assume that the functional form of the portfolio policy is correct, to replace $w_{i,t}$ with a function of $x_{i,t}$, and that the coefficients are constant through time, to condition down the conditional expectation. If either assumption is incorrect, the marginal utilities in equation (9) will be correlated with variables in the investor's information set at date t, which may include missing characteristics or variables that are correlated with the variation in the coefficients. We can therefore perform specification tests for the portfolio policy using the standard overidentifying-restrictions test of Hansen (1982).

Finally, note that the method of moments interpretation does not necessarily render our approach frequentist and therefore unable to accommodate finite-sample uncertainty about the parameters and model specification. Nigmatullin (2003) shows how to interpret first-order conditions similar to equation (9) from a Bayesian perspective using the idea of an empirical likelihood function and explains how to incorporate parameter and model uncertainty. While his application deals with the nonparametric approach of Ait-Sahalia and Brandt (2001), the general idea applies directly to our approach.

2.3 Refinements and Extensions

Besides its effectiveness and simplicity, an important strength of our approach is that the basic idea is easily refined and extended to suit specific applications. We now discuss some of the possible refinements and extensions to illustrate the flexibility of our approach.

2.3.1 Objective Functions

The most important ingredient of any portfolio choice problem is the investor's objective function. In contrast to the traditional Markowitz approach, our specification of the portfolio choice problem can accommodate any choice of objective function. The only implicit assumption is that the conditional expected utility maximization problem (1) be well specified with a unique solution. Besides the standard HARA preferences (which nest constant relative risk aversion, constant absolute risk aversion, log, and quadratic utility), our approach can also be applied to behaviorally motivated utility functions, such as loss aversion, ambiguity aversion, or disappointment aversion, as well as practitioner-oriented objective functions, including maximizing the Sharpe or information ratios, beating or tracking a

benchmark, controlling draw-downs, or maintaining a certain value-at-risk (VaR).[7]

In the empirical application we use standard constant relative risk aversion (CRRA) preferences over wealth:

$$u(r_{p,t+1}) = \frac{(1 + r_{p,t+1})^{1-\gamma}}{1 - \gamma}. \tag{13}$$

The advantage of CRRA utility is that it incorporates preferences toward higher-order moments without introducing additional preference parameters. In addition, the utility function is twice continuously differentiable, which allows us to use more efficient numerical optimization algorithms that make use of the analytic gradient and Hessian of the objective function. We also offer results for the minimum variance and maximum Sharpe ratio portfolios.

2.3.2 Portfolio Weight Constraints

By far the most common departure from the basic portfolio choice problem (1) in practice are constraints on the optimal portfolio weights. In our approach, these constraints have to be imposed through the parameterization of the portfolio policy. For example, consider the case of the no-short-sale constraint in long-only equity portfolios. The simplest way to impose this constraint through the portfolio policy is to truncate the portfolio weights in equation (3) at zero. Unfortunately, in doing so the optimal portfolio weights no longer sum to one (setting the negative weights to zero results in an sum of weights greater than one). We therefore need to renormalize the portfolio weights as follows:

$$w_{i,t}^+ = \frac{\max[0, w_{i,t}]}{\sum_{j=1}^{N_t} \max[0, w_{j,t}]}. \tag{14}$$

Besides guaranteeing positivity of the portfolio weights, this specification is also an example of a nonlinear parameterization of the portfolio weight function (2).

One computational problem with this specification of the portfolio policy function is its non-differentiability at $w_{i,t} = 0$. In order to compute the standard errors of the estimated θ from first-order conditions analogous to equation (9), we require first-order derivatives. One

---

[7]Benartzi and Thaler (1995), Ait-Sahalia and Brandt (2001), Ang, Bekaert, and Liu (2003), Gomes (2003), among other, examine the role of behaviorally motivated preference in portfolio choice. Practitioner oriented objective functions are considered, for example, by Roy (1952), Grossman and Vila (1989), Browne (1999), Tepla (2001), Basak and Shapiro (2001), and Alexander and Baptista (2002).

way to overcome this problem in practice is to approximate the function $\max[0, y]$ between two close points $y = 0$ and $y = \alpha > 0$ with either a third or a fifth-order polynomial with smooth first- or first- and second-order derivatives at the end points, respectively.

2.3.3 Nonlinearities and Interactions

Although we explicitly specified the portfolio policy as a linear function of the characteristics, the linearity assumption is actually innocuous because the characteristics $x_{i,t}$ can always contain nonlinear transformations of a more basic set of characteristics $y_{i,t}$. This means that the linear portfolio weights can be interpreted as a more general portfolio policy function $w_{i,t} = \bar{w}_{i,t} + g(y_{i,t}; \theta)$ for any $g(\cdot; \cdot)$ that can be spanned by a polynomial expansion in the more basic state variables $y_{i,t}$. Our approach therefore accommodates very general departure of the optimal portfolio weights from the benchmark weights.

Cross-products of the characteristics are a particularly interesting form of non-linearity because they have the potential to capture interactions between the characteristics. For instance, there is considerable evidence in the literature that the momentum effect is concentrated in the group of growth (low book-to-market) firms (e.g., Daniel and Titman, 1999). Our approach can capture this empirical regularity by including the product of the book-to-market ratio and the one-year lagged return as an additional characteristic.

In practice, we need to choose a finite set of characteristics as well as possible nonlinear transformations and interactions of these characteristics to include in the portfolio policy specification. This variable selection for modeling portfolio weights is no different from variable selection for modeling expected returns with regressions. The characteristics and their transformations can be chosen on the basis of individual t tests and joint F tests computed using the covariance matrix of the coefficient estimates, or on the basis of out-of-sample performance.

2.3.4 Time-Varying Coefficients

The critical assumption required for conditioning down the expectation to rewrite the conditional problem (1) as the unconditional problem (4), is that the coefficients of the portfolio policy are constant through time. While this is a convenient assumption, there is no obvious economic reason for the relation between firm characteristics and the joint

distribution of returns to be time-invariant. In fact, there is substantial evidence that economic variables related to the business cycle forecast aggregate stock and bond returns.[8] Moreover, the cross-section of expected returns appears to be time-varying as a function of the same predictors (e.g., Cooper, Gulen, and Vassalou, 2000).

To accommodate possible time-variation in the coefficients of the portfolio policy, we can explicitly model the coefficients as functions of the business cycle variables. Given a vector of predictors observable at date t, denoted by z_t, we specify the portfolio policy has:

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t} \theta^\top (z_t \otimes x_{i,t}) \tag{15}$$

where $\otimes$ denotes the Kronecker product of two vectors. In this form, the impact of the characteristics on the portfolio weight varies with the realization of the predictors z_t.

2.3.5 Multi-Period Horizon

The discussion thus far assumed a myopic single-period horizon. However, our approach can be extended to a multi-period horizon problem as follows. Consider an investor who maximizes the utility of a K-period return:

$$r_{p,t+1 \to t+K} = \prod_{k=0}^{K-1} (1 + r_{p,t+k+1}) - 1, \tag{16}$$

where each period's return is given as before by:

$$
\begin{aligned}
r_{p,t+k+1} &= \sum_{i=1}^{N_{t+k}} w_{i,t+k} r_{i,t+k+1} \\
&= \sum_{i=1}^{N_{t+k}} \left(\bar{w}_{i,t+k} + \frac{1}{N_{t+k}} \theta_k^\top x_{i,t+k} \right) r_{i,t+k+1}
\end{aligned} \tag{17}
$$

with period-specific portfolio policy coefficients θ_k. Following the same steps as in the single-period problem, the investor's conditional expected utility optimization with respect to the portfolio weights $\{\{w_{i,t+k}\}_{i=1}^{N_{t+k}}\}_{k=0}^{K-1}$ can be alternatively expressed as an unconditional

---

[8] For example, Campbell (1991), Campbell and Shiller (1988), Fama (1990), Fama and French (1988,1989), Hodrick (1992), and Keim and Stambaugh (1986) report evidence that the stock market returns can be forecasted by the dividend-price ratio, the short-term interest rate, the term spread, and the credit spread.

expected utility optimization with respect to the portfolio policy coefficients $\{\theta_k\}_{k=0}^{K-1}$. The sample analogue of this unconditional problem is:

$$\max_{\{\theta\}_{k=0}^{K-1}} \frac{1}{T} \sum_{t=0}^{T-1} u \left(\prod_{k=0}^{K} r_{p,t+1 \to t+K} \right). \tag{18}$$

Expressing the investor's portfolio choice problem in terms of optimizing over portfolio policies, as opposed to over conditional portfolio weights, has the same conceptual benefits in the multi-period problem as we discussed in the single-period context (e.g., it avoids modeling the joint distribution of returns and characteristics). In addition, however, it allows for the universe of investable securities to change over the investor's horizon, meaning that N_{t+k} can be a random variable for $k > 0$. Incorporating this realistic feature of the portfolio problem in the traditional approach is difficult, if not practically impossible, because it would require an explicit model of exchange listings and delistings.

Although theoretically sound, there are at least two practical complications introduced by having a multi-period horizon. First, the number of coefficients increases with the number of periods, raising the risk of in-sample overfitting. This problem can be overcome by imposing a parametric structure on the coefficients of the portfolio policy across periods, such as the restriction that the coefficients are constant or that the coefficients for each period are scaled by some function of the number of periods to the end of the horizon. The second complication is that if the average utility in equation (18) is computed with overlapping observations, the time-dependence that is induced by the overlap can severely distort the statistical inferences based on the iid covariance matrix estimator (12). While it is possible to correct the covariance matrix estimator for the persistence introduced by the overlapping observations, the resulting inferences are still questionable (e.g., Valkanov (2003)). The best way to avoid this problem is to use only T/K non-overlapping observations in constructing the sample analogue of the unconditional optimization problem. Doing so unfortunately limits the length of the investment horizon we can realistically consider with our approach given the finite size of the sample. In the empirical application, we use overlapping observations with investment horizons up to five years and compute the standard errors using bootstrap in order to obtain correctly-sized, albeit conservative, tests.

13

2.4　Testing Asset Pricing Models

Portfolio choice problems are by nature normative, prescribing what the investor should do to maximize expected utility. An interesting feature of our approach is that it also has a positive application, helping economists understand better how markets function. Specifically, the cross-sectional relation of expected returns in an equilibrium asset pricing model is typically derived by combining the first-order conditions of the portfolio choice problem of a representative agent with the market clearing condition that this agent has to hold each asset in proportion to its market capitalization weight.

Reversing this argument, a given preference schedule for the representative agent is only consistent with market equilibrium if the agent's optimal portfolio weights equal the observed market cap weights. Our approach is ideally suited for testing this implication, since we model explicitly the deviations of the investor's optimal portfolio weights from the observed market cap weights. A rejection of the hypothesis $\theta = 0$ implies that the preferences do not correspond to those of the representative agent.

To be more specific, any equilibrium asset pricing model (CAPM, equilibrium APT, or CCAPM), requires that the representative investor hold all the assets in the economy in the proportions they exist in the market. Our proposed test makes an assumption about the utility function of the representative investor (e.g., CRRA) and about the functional form of the portfolio policy but does not require that we model the expected returns and risk of stocks (e.g., measuring risk as covariance with the market or covariance with consumption). We just need to find the agent's optimal portfolio and compare it to the market portfolio.[9]

Consumption-based asset pricing models can be tested within our approach by positing a representative investor with utility for consumption over time. The investor's consumption policy can be parameterized along the same lines and in addition to the parameterized portfolio policy. The coefficients of both policy functions can be estimated by maximizing utility in sample. We can then compare the utility obtained by the investor that follows these optimal rules with the utility obtained from the observed consumption flow. This is similar to the approach of Brandt (1999).

Furthermore, our tests can easily accommodate recent developments in the asset pricing literature that justify the existence of equilibrium priced risk factors in the stock

---

[9]Actually, to the extent that the weight function is misspecified, that makes the test conservative.

market. In addition to investment risks, the investor may face background risks related to human capital (Jagannathan and Wang, 1996) or private equity (Heaton and Lucas, 2000). The correlation of risk factors other than the market with such background risks explains why the investor chooses to hold the market in equilibrium and does not tilt the portfolio towards the risk factors. Background risks can be easily included as additional arguments in the utility function.

Merton (1973) showed that time variation in consumption or investment opportunities may induce portfolio hedging demands. This may explain the risk premia offered by assets that are correlated with the changes in the opportunity set. Campbell and Vuolteenaho (2004), Brennan and Xia (2004), and Bansal, Dittmar, and Lundblad (2004) justify the premium of value stocks relative to growth stocks by their poor performance at times when the market premium is low. According to such explanations, long-lived investors do not want to deviate from the market portfolio towards the high-return stocks because they find them riskier in this intertemporal sense. These explanations can be tested by estimating the optimal portfolio of a multi-period horizon investor as explained in section 2.3.5, possibly using portfolio policies that depend on macro state variables as in section 2.3.4.

3   Empirical Application

To illustrate the simplicity, the flexibility, and, most importantly, the effectiveness of our approach, we present an empirical application involving the universe of all listed stocks in the U.S. from January of 1964 through December of 2002. We first describe the data and then present results for the base case, various extensions, and an out-of-sample experiment. Unless otherwise stated, we assume an investor with CRRA preference and a relative risk aversion of five. In the application, the investor is restricted to only invest in stocks. We do not include the risk-free asset in the investment opportunity set. The reason is that the first-order effect of allowing investments in the risk-free asset is to vary the leverage of the portfolio, which only corresponds to a change in the scale of the stock portfolio weights.

3.1 Data

We use monthly firm-level returns from CRSP as well as firm-level characteristics, obtained from the CRSP-Compustat merged dataset, from January of 1964 through December of 2002. For each firm in the CRSP-Compustat dataset, we construct the following variables at the end of each fiscal year: the log of the firm's market equity (me), defined as the log of the price per share times the number of shares outstanding, and the firm's log book-to-market ratio (btm), defined as the log of one plus book equity (total assets minus liabilities, plus balance-sheet deferred taxes and investment tax credits, minus preferred stock value) divided by market equity.[10] We use the standard timing convention of leaving at least a six-month lag between the fiscal year-end characteristics and the monthly returns, to ensure that the information from the annual reports would have been publicly available at the time of the investment decision. From the CRSP database, we record for each firm the lagged one-year return (mom) defined as the compounded return between months $t - 13$ and $t - 2$. Similar definitions of the three characteristics are commonly used in the literature (e.g., Fama and French, 1996). The Appendix provides further details about the firm-level data, including the exact definitions of the components of each variable. We use size, book-to-market, and momentum as conditioning characteristics in the portfolio optimization since we want to compare our results with previous studies and these characterists are the most widely used in the literature.

The number of firms in our sample is generally trending upward, with an average annual growth rate of 4.2%. The average number of firms throughout our sample is 3,680, with the fewest firms in February of 1964 (1,033 firms) and the most firms in November of 1997 (6,356 firms).

Figure 1 describes the three firm characteristics. The first column plots the cross-sectional means of the (non-standardized) characteristics at each month in our sample. The second column shows the corresponding cross-sectional standard deviations. Recall that the characteristics enter the portfolio policy function in standardized form. The plots in Figure 1 can be used to translate given values of the standardized characteristics at a particular date in the sample into the original characteristics at the same date.

We use the one-month Treasury bill rate as the riskfree rate. In an extension of our

---

[10]Taking logs makes the cross-section distribution of me and btm more symmetric and reduces the effect of outliers.

basic approach, we model the coefficients of the portfolio policy as functions of the term spread (tsp), defined as the difference in the yields to maturity of the 10-year Treasury note and the one-year Treasury bill. Monthly interest rate data is obtained from the DRI database for the same sample period as the stock data.

3.2   Base Case

Table 1 presents the results for the base case, in which the over- or under-weighting of each stock, relative to the value-weighted market portfolio, depends on the firm's market capitalization, book-to-market ratio, and lagged one-year return, using the policy function in equation (3). The table is divided into four sections describing separately the (i) parameter estimates and standard errors, (ii) distribution of the portfolio weights, (iii) properties of the optimized portfolio returns, and (iv) average characteristics of the portfolio. This format is the same for all tables in the paper. In addition to the results for the universe of all CRSP stocks, we also present results for a restricted set of the 500 largest firms. We first focus on the results for all stocks and then point out where the results for the top 500 stocks differ.

The first few rows in Table 1 present the estimated coefficients of the portfolio policy along with their standard errors estimated from 1,000 bootstrapped samples.[11] For the universe of all stocks, in the second column, the deviations of the optimal weights from the benchmark weights decrease with the firms's market capitalization (size) and increase with both the firm's book-to-market ratio (value) and its lagged one-year return (momentum). The signs of the estimates are consistent with the literature. The investor over-weights small firms, value firms, and past winners and under-weights large firms, growth firms, and past losers. Since the characteristics are standardized cross-sectionally, the magnitudes of the coefficients can be compared to each other. Quantitatively, a high book-to-market ratio leads to the largest over-weighting of a stock. All three coefficients are highly significant. We also test whether all three coefficients are jointly equal to zero using a Wald test, and the bootstrapped p-value of this test is reported in the row labeled "Wald p-value."[12]

The next few rows describe the weights of the optimized portfolio (in the second column) and compare them to the weights of the market portfolio (in the first column). The

---

[11]We use bootstrapped standard errors since they produce slightly more conservative tests (larger standard errors) than using estimates of the asymptotic covariance matrix in equation (10).

[12]When the bootstrapped p-value from the Wald test is less than 0.001, we report it as 0.000.

average absolute weight of the optimal portfolio is about three times that of the market (0.12% versus 0.04%). Not surprisingly, the active portfolio takes larger positions. However, these positions are not extreme. The average (over time) maximum and minimum weight of the optimal portfolio are 4.36% and -0.32%, respectively, while the corresponding extremes for the market portfolio are 4.63% and 0.00%. The average sum of negative weights in the optimal portfolio is -124%, which implies that the sum of long positions is on average 224%. Finally, the average fraction of negative weights (shorted stocks) in the optimal portfolio is 0.468. Overall, the optimal portfolio does not reflect unreasonably extreme bets on individual stocks and could well be implemented by a combination of an index fund that reflects the market and a long-short equity hedge fund.

The following rows characterize the performance of the optimal portfolio relative to the market portfolio. For ease of interpretation, all measures are annualized. The optimal portfolio has a volatility slightly larger than that of the market portfolio, 19% versus 16%, but has a much higher average return of 24.4% as opposed to 12.0% for the market. This translates into a Sharpe ratio that is three times the market's and a certainty equivalent gain of 10%. We can use a regression of the excess returns of the active portfolio on the excess return of the market to evaluate the active portfolio's alpha, market beta, and residual risk, and then use these statistics to compute the portfolio's information ratio. The alpha of the portfolio is over 15%, with a low market beta of only 0.440. Dividing the alpha by the residual volatility of 17.7% produces an information ratio of 0.890. We should point out that it is not surprising that the optimal portfolio outperforms the market, because we have chosen characteristics that are known to be associated with substantial risk-adjusted returns.

We can decompose of the optimal portfolio returns into the market return and the return on a long-short equity hedge fund along the lines of equation (8). The average return of this hedge fund is found to be 12.46% (not shown in the table). We can further decompose the hedge fund return as $r_h = q(r_h^+ - r_h^-)$ where r_h^+ is the return on the long part of the hedge fund and r_h^- is the return on the short part, both normalized such that the sum of their weights is one. In this way, q captures the leverage of the long-short portfolio. The average r_h^+ is 19.00% and the average r_h^- is 11.84%, so that the return of the hedge fund without leverage, i.e., with one dollar long and one dollar short positions, is 7.16%. These returns compares with the market's return of 11.96% over the same period. We therefore see that the long side of the hedge outperforms the market whereas the short side has roughly

the same performance as the market. In fact, the short side could be replaced with a short position in the market portfolio without hurting performance. This is important since it is obviously easier to short the market using futures than it is to hold a short portfolio of stocks. The average return of the entire hedge fund of 12.46% and the returns of the scaled long and short parts imply a leverage q of the long and short positions of the order of 170%.

To describe the composition of the optimized portfolio, we compute for every month the weighted characteristics of the portfolio as $N_t \sum_{i=1}^{N_t} w_{i,t}\hat{x}_{i,t}$. The last three rows of the table compare the average (through time) weighted characteristics of the optimized portfolio to those of the market portfolio. The market portfolio has a bias toward very large firms (due to value weighting) and firms with below-average book-to-market ratios (growth), while it is neutral with respect to momentum. In contrast, the optimized portfolio has a slight bias toward small firms and much stronger biases toward high book-to-market ratio (value) firms and past winners. Specifically, the portfolio's book-to-market ratio is more than three standard deviations above the average stock, and the portfolio's momentum is close to two standard deviations above the mean.[13]

Figure 2 plots the time series of the three portfolio characteristics. The characteristics vary over time, but their variability is relatively small and they appear stationary. Moreover, the book-to-market characteristic is always larger than the momentum characteristic, which in turn is larger than the size characteristic, indicating that the optimized portfolio reflects consistent bets through time. While this ordering is also clearly captured in the averages reported in the table, it is comforting to note that the results are systematic and not the product of a few outliers. Finally, one might suspect that maintaining such stable portfolio characteristics requires unreasonably large trading activity. Fortunately, this is not the case. The average turnover of the optimized portfolio is only 50% per year, as compared to a average turnover of 12% per year for the market portfolio (due to new listings, delistings, equity issues, etc). This further shows that the optimal portfolio is eminently implementable and that the returns are unlikely to be affected much by trading costs.

The results are similar for the 500 largest stocks, which are an interesting subset of firms because they are certainly liquid enough to implement an active portfolio strategy in large scale. The most notable difference in the results is that the book-to-market coefficient,

---

[13]In a long-short portfolio, this does not necessarily mean that the typical stock has characteristic values of this order of magnitude. For instance, a portfolio that is long 200% in stocks with a characteristic value of 2 and is short 100% of stocks with a characteristic value of 1 has an average characteristic value of 3.

while still positive and significant, is about half as large as in the case of all stocks. The optimal allocation is achieved with a lower average number of stocks shorted, 35.8%, and the total short position of the optimal portfolio is only -51.8%. The average return of the optimal portfolio is 20%, or more than 7% larger than the return of the market portfolio. Its standard deviation, at 17.6%, is only slightly larger than that of the benchmark. These combine into a Sharpe ratio that is more than twice the market's. The alpha of the optimal portfolio is 10% with a market beta of 0.704, residual volatility of 13.6% and a information ratio of 0.734. The average characteristics of the optimized portfolio are again tilted toward smaller, value, and winner stocks.

Finally, we can test equilibrium in the stock market for a representative investor with these preferences. The Wald test of the joint significance of the portfolio policy coefficients has a p-value of less than 1% for both the entire universe of stocks and the subset of the largest 500 stocks. This evidence goes against Fama and French (1993, 1996) who argue that the association between firm characteristics and stock returns arises because the characteristics proxy for exposure to factor risk. If indeed the premia associated with the characteristics were only fair compensation for their risk, the investor would choose to hold the market instead of tilting the portfolio policy towards the characteristics. We strongly reject that hypothesis. In contrast, our findings support Daniel and Titman (1997), who argue that the characteristics do explain expected returns but are not associated with a corresponding exposure to systematic risk.

3.3 Extensions

3.3.1 Portfolio Weight Constraints

A large majority of equity portfolio managers face short-sale constraints. In Table 2, we present the results from estimating the long-only portfolio policy specified in equation (14). As in the unconstrained case, the deviation of the optimal weight from the market portfolio weight decreases with the firm's size, increases with its book-to-market ratio, and increases with its one-year lagged return. Focusing on the portfolio involving the entire universe of stocks, a high book-to-market ratio and large positive one-year lagged return are less desirable characteristics for a long-only investor. The coefficients associated with both of these characteristics are lower in magnitude than in the unrestricted case and are only

marginally significant, whereas the coefficient associated with the market capitalization of the firm is not significant. Overall, the significance of the θ coefficients is substantially diminished compared to the unconstrained base case.

The optimal portfolio still does not involve extreme weights. In fact, the average maximum weight of the optimal portfolio is only 2.5%, which is actually lower than that of the market portfolio. On average, the optimal portfolio invests in only 56% of the stocks. The resulting mean and standard deviation of the portfolio return are 17.3% and 10.0%, respectively, translating into a certainly equivalent gain of 3.3% relative to holding the market portfolio. The alpha, beta, and information ratio of the portfolio are 5.4%, 0.977, and 0.620, respectively. These statistics are quite remarkable, given the long-only constraint. The average size of the firms in the optimal portfolio is greater than the size of the average firm but significantly lower than that of the value-weighted market portfolio. The book-to-market ratio and momentum characteristics are less than one standard deviation above those of the average stock and are also significantly different from those of the market portfolio. The results for the optimal long-only portfolio in the universe of the top 500 stocks are qualitatively similar.

The most interesting comparison is between the long-only portfolio in Table 2 and the unconstrained base case in Table 1. Qualitatively, the results in the two tables are similar, but the long-only constraint results in a lower average portfolio return, slightly higher volatility, and consequently a lower certainty equivalent gain relative to holding the market portfolio. The difference in performance is due to two related factors. First, the unconstrained portfolio can exploit both positive and negative forecasts, while the constrained portfolio can only exploit the positive forecasts. Consistent with this argument, the fraction of short positions in Table 1 is roughly the same as the fraction of stocks not held by the long-portfolio in Table 2. Second, the unconstrained portfolio benefits from using the short positions as leverage to increase the exposure to the long positions.

Interestingly, the tests for joint significance of all three parameters have a p-value slightly above 10%. We therefore cannot reject that the coefficients are jointly zero and that the investor is equally well off holding the market as holding the optimal portfolio. This rejection is consistent with the increase in the standard errors on the coefficients and the smaller gain in certainty equivalent of the restricted optimal portfolio relative to the market. However, we should note that part of this rejection is due to the conservative nature of the Wald test, which strongly penalizes the insignificance of the coefficient on market equity even

though the *t*-tests found the coefficients on book to market and momentum to be individually significant. Despite this caveat, taking the test at face value has the interesting implication that we no longer reject equilibrium in the stock market. Now, our tests side with Fama and French (1993, 1996) and contradict Daniel and Titman (1997). We conclude that short sales constraints go a long way towards explaining the size, value, and momentum anomalies.

3.3.2   Interactions

Table 3 includes interactions between the three characteristics in both the constrained and unconstrained portfolio policies. Specifically, we include the levels and the cross-products of the firms' market capitalization, book-to-market ratio, and one-year lagged return, resulting in a total of six characteristics. The interactions do not substantially change the essence of the previous results. Both portfolios are still tilted toward smaller firms, value firms, and past winners.

Examining more closely the estimated coefficients for the interacted characteristics, all three coefficients are statistically significant in the unconstrained case. Relative to the base case, the optimal portfolio policy tilts the weights positively towards large-growth and small-value stocks, large-winner and small-loser stocks, and value-winner and growth-loser stocks. The policy tilts the weights negatively towards the opposite combinations of characteristics. It is important to notice that the coefficients can only be interpreted as marginal contributions to the weights (evaluated at the average level of characteristics), much like the coefficients of multivariate regressions cannot be interpreted in isolation. To understand the result of the optimal policy, it is best to examine the resulting average characteristics in the portfolio. The average characteristics are comparable to the base case in Table 1. However, a substantial portion of the book-to-market exposure is now obtained in smaller firms and past winners. In the constrained case, the coefficients on the interaction terms are generally smaller and less significant. The size-momentum interaction does not play a significant role in the constrained portfolio, whereas the value-momentum interaction is still important.

The interactions do not lead to dramatic changes in the distribution of the optimal portfolio weights, but the effect on the distribution of the portfolio return is substantial. Both the unconstrained and constrained portfolios exhibit higher average returns (42.8% and 20.0%), higher Sharpe ratios (1.318 and 0.746) and certainly equivalents (23.1% and

10.6%), higher alphas (32.4% and 8.1%), and higher information ratios (1.283 and 0.795). Comparing these numbers to the corresponding results without interactions, in Tables 1 and 2, the performance improvement is much greater for the unconstrained portfolio. The certainty equivalent increases by 7.8% in the unconstrained case and by only 1.4% in the constrained case.

Taking interactions between the characteristics into account leads to a sound rejection of stock market equilibrium with this representative agent. The p-values of the Wald test are now indistinguishable from zero both for the long-short and the long-only optimal portfolios.

Besides illustrating the potential for interactions between characteristics to improve the performance of the optimized portfolio, especially in the unconstrained case, this extension also emphasizes the ability of our method to accommodate a larger number of characteristics. Given analytical expressions for the gradient and Hessian of the objective function, the estimation with six characteristics takes only marginally longer than in the base case with three characteristics.

3.3.3 Time-Varying Coefficients

In Table 4 we allow the coefficients of the portfolio policy to depend on the slope of the yield curve. We estimate a separate set of coefficients for months when the yield curve at the beginning of the month is positively sloped (normal) and negatively sloped (inverted). Since inverted yield curves tend to be associated with recessions, letting the portfolio coefficients vary with the yield-curve slope allows the effect of the characteristics on the joint distribution of returns to be different during expansionary and contractionary periods.

The table presents results for unconstrained and constrained portfolios. In both cases, the most dramatic effect of conditioning on the slope of the yield curve is on the role of the firm's size. When the yield curve is upward sloping, the optimal portfolio is tilted toward smaller firms, just as in the base case. When the yield curve is downward sloping, in contrast, the tilt is exactly the opposite, with a positive coefficient (although not statistically different from zero in the unconstrained case). This is consistent with the common notion that small firms are more affected by economic downturns than larger and more diversified firms. For book-to-market and momentum, the coefficients are generally larger in magnitude when the yield curve slopes down.

23

Conditioning on the slope of the yield curve does not significantly alter the distribution of the optimal portfolio weights. However, the performance of the portfolio is improved. Both portfolios have higher average returns, certainly equivalents, alphas, and information ratios than without conditioning. The certainty equivalent improvement relative to the base case is again greater for the unconstrained portfolio, with a gain of 1.8% versus 0.2% for the constrained portfolio.

The average characteristics of the optimal portfolios are the most interesting to analyze. Consider first the unconstrained case. As suggested by the coefficient estimates, the optimal portfolio is tilted toward small stocks when the yield curve is upward sloping. When the yield curve is downward sloping, the portfolio is tilted toward larger stocks and resembles closely the composition of the market portfolio. The average book-to-market and momentum characteristics are both positive and larger when the yield slope is positive. Turning to the constrained results, notice that the optimal portfolio is not significantly tilted into value and winner firms when the yield curve is inverted, suggesting that during those times the profitability of value and momentum comes primarily from short positions.

When we condition the portfolio policy on the slope of the yield curve, the Wald test rejects at the 1% level the hypothesis that the market is the optimal portfolio, both for the unconstrained and the positive-only portfolios. Since the slope of the yield curve is associated with the business cycle, our results do not support explanations of the value premium based on macroeconomic risk stories. We find that the investor actually chooses to invest more in value stocks during recessions.

3.3.4 Objective Functions

The optimal portfolio policy depends critically on the investor's preferences. The results thus far were obtained assuming CRRA utility with relative risk aversion $\gamma = 5$. To get a better sense for the role of this utility assumption, we present in Table 5 results for different levels of risk aversion. In addition to $\gamma = 5$, which we report in the table for comparison, we also estimate the optimal portfolio for $\gamma = 1$, corresponding to the popular case of log-utility, and $\gamma = 100$, which is extremely high and makes the investor very sensitive to extreme losses.

The table presents results for both constrained and unconstrained portfolios, but the discussion is focused on the unconstrained case and only mentions the constrained case when there are meaningful differences. For small values of γ, the estimates of the coefficients on

the firm's size, book-to-market ratio, and one-year lagged return are all large in absolute value and statistically significant. As the investor becomes more risk averse, the coefficients on size and momentum approach zero. This suggests that these characteristics are associated with both mean returns and risk. As risk aversion increases, the investor weighs more the contribution of these characteristics to risk and loads less heavily on them. In contrast, the exposure to book-to-market does not change qualitatively as risk aversion increases. This indicates that this characteristic is more associated with expected return than risk.

The average firm characteristics exhibit the same patterns. For $\gamma = 1$, the portfolio is severely tilted toward firms that are small, value, and winners. As the level of risk aversion increases, the tilting towards small caps and winners decreases. Actually, for $\gamma = 100$, the portfolio holds companies that are one standard deviation larger than the mean. However, the tilt towards value firms is maintained.

The distribution of the optimal portfolio weights also changes with the level of risk aversion. In particular, an investor with $\gamma = 1$ takes on more and larger negative positions, compared to an investor with $\gamma = 5$. The fraction of shorted stocks is only increased by 5% but the sum of negative weights is three times larger, which implies that the less risk averse investor takes similar bets but with more leverage. Interestingly, the $\gamma = 100$ investor actually uses higher leverage than the investor with $\gamma = 5$. Intuitively, the short positions help by partially hedging the worst performing stocks in the market's lowest return months.

Not surprisingly, the differences in the optimal portfolio weights translate into equally striking differences in the distribution of the optimized portfolio returns. The average return and volatility are highest for $\gamma = 1$ and decrease with the investor's level of risk aversion. Most importantly, the portfolio's minimum return (not shown in the table) decreases from -67.20% for $\gamma = 1$, to -19.77% for $\gamma = 5$, and to -13.19% for $\gamma = 100$ (for comparison, the market's minimum return is -21.50%). For high levels of γ, the curvature of the utility function is such that the average utility across all months is dominated by the utility obtained in the worst month. In this sense, the $\gamma = 100$ preferences correspond closely to a max-min criterion. The table also presents certainty equivalents for the different levels of risk aversion but they cannot be compared with each other. The certainty equivalent for the $\gamma = 100$ case is actually negative, -0.976, but the investor dislikes the market even more, with a certainty equivalent of -1.976.[14]

---

[14]The certainty equivalents can take values less than -1 because we are taking the monthly certainty equivalents and multiplying them by 12 to express them in annual terms. Of course, the monthly certainty

The Wald tests reject equilibrium in the stock market for all the unrestricted portfolios and also for the restricted portfolio when the investor has log utility. Although increasing risk aversion helps in explaining the size and momentum anomalies it does not explain the value anomaly.

As another example of how the objective function affects the optimal portfolio policy, Table 6 presents results for an investor who wants to minimize the variance or maximize the Sharpe ratio of the portfolio. This application is particularly interesting in that the resulting two portfolios span the mean-variance frontier.

We find that the minimum variance portfolio only loads significantly on the book-to-market characteristic. This result further deepens the value anomaly. Even when the investor does not care about expected returns but only about risk, the portfolio still loads on stocks with high book-to-market. In the long-only case, the coefficient are generally insignificant, resulting in average characteristics that approximate the market's.

On average, short positions constitute 52.0% of the portfolio, substantially less than in the base case. The minimum variance portfolio therefore makes relatively little use of leverage. The standard deviation achieved by this portfolio is 13.7%. This is substantially higher than the minimum variance of 6.85% per year obtained by Chan, Karceski, and Lakonishok (1999) using the realized covariance matrix of returns to find the minimum variance portfolio. The difference highlights the relatively poor performance of the size, book-to-market, and momentum characteristics in capturing the risk of equity portfolios.

The maximum Sharpe ratio portfolio loads significantly on the three characteristics in the long-short case and on book-to-market and momentum in the long-only case. This results in high average book-to-market and momentum for the unconstrained portfolio whereas, for the long portfolio, only momentum is more than one standard deviation higher than the mean. Interestingly, the exposures of the maximum Sharpe ratio portfolio closely resemble the optimal portfolio for the $\gamma = 5$ CRRA investor (Tables 1 and 2). In other words, the optimal investment policy of an investor who maximizes the Sharpe ratio is similar to that of a CRRA investor with $\gamma = 5$. The maximized Sharpe ratio is almost three times the market's in the unconstrained case and two times the market's in the long-only case.

equivalents cannot be less that -1 no matter the degree of risk aversion.

26

3.3.5 Multi-Period Horizon

Finally, we present an application of our approach to a multi-period problem. Table 7 shows the optimal portfolio policy of an investor with a multi-period horizon of one or five years (the one-month results are reproduced for comparison) with monthly rebalancing. The portfolio policies are estimated from the maximization problem (18) using overlapping returns and the standard errors are obtained via bootstrap.

The optimal portfolio policy changes little when the investment horizon goes from one month to one year. However, for a five-year horizon, the loading on momentum almost doubles. As a consequence, the leverage of the optimal portfolio also increases substantially. Note that the Sharpe ratios of the portfolios are quite similar. However, the (annualized) certainty equivalent of the five-year horizon investor is two times higher than for the shorter-horizon portfolios. This happens because the five-year returns are actually less risky than the standard deviation (computed from monthly returns) indicates. In fact, the average portfolio return in the sample is so high that there is no negative realization in any five-year period. However, this only happens for the long-short portfolio. The certainty equivalent of the long-only investor does not change substantially with the increasing horizon.

It is striking that the multi-period investor still finds it optimal (in fact, even more so) to tilt the portfolio toward value and momentum as the horizon increases. This contrasts with Brennan and Xia (2004), Campbell and Vuolteenaho (2004), and Bansal, Dittmar, and Lundblad (2004), who argue that the value premium can be explained as an ICAPM risk premium. Their argument is that value stocks offer bad returns when investment opportunities are poor. Value stocks would therefore be required to offer a higher expected return to an investor with a long horizon who cares about reinvestment risk. However, we find the opposite result. A five-year horizon investor actually wants to take a larger bet on value stocks. Admitedly, five years is not a very long horizon and, in particular, is quite short if time is measured in business cycles. On the other hand, Barberis (2000) and Brandt (1999) show that an increase in horizon from five to, say, ten years would typically have little impact on portfolio choices. Moreover, our tests are quite conservative in that the portfolio policy forces the investor to hold the same portfolio weights through time. This is restrictive since, in general, the optimal policy would be dependent on the time to horizon. The fact that, even in this suboptimal case, the investor deviates significantly from the market makes the tests even more powerful. Similar to value, the exposure to momentum also increases with

the horizon, suggesting that the momentum premium cannot be justified as compensation for intertemporal risk. Finally, we again find that, with short-sale constraints, we cannot reject the optimality of the market portfolio.

3.4 Out-of-Sample Performance

While the stellar performance of our approach is unlikely to be due to over-fitting since we optimize a portfolio with a large number of stocks over a small number of parameters, the most convincing way to establish its robustness is through an out-of-sample experiment. We split our sample into two equal sub-samples from January of 1964 through June of 1983, and from July of 1983 through December of 2002, labeled Sample 1 and Sample 2, respectively. For each sub-sample, we estimate the optimal portfolio policy as in the base case and then use the estimated coefficients to form a portfolio in the other sub-sample. More specifically, for each month in Sample 2, we form a portfolio using the firm characteristics observed at the beginning of that month and the policy coefficients estimated from Sample 1. The returns generated by these portfolios in Sample 2 represent the out-of-sample results for Sample 1. We conduct an analogous out-of-sample experiment for Sample 2, flipping the estimation and evaluation samples from the previous case.

Panel A of Table 8 presents the unconstrained results. Before discussing the out-of-sample performance of our approach, we briefly comment on the sub-sample results. The coefficient estimates are generally consistent across sub-samples. The size coefficient is negative and the book-to-market and momentum coefficients are positive in both sub-samples. The coefficients are individually and jointly significant even with the smaller sample sizes, which implies that our tests are quite powerful. The distributions of the weights and returns of the optimized portfolio are also similar across sub-samples. The certainty equivalent is 15.1% in the first and 17.7% in the second sub-sample. Finally, in both sub-samples, the portfolios are tilted substantially toward value and momentum, with momentum playing a larger role in the first sub-sample and book-to-market playing a larger role in the second sub-sample.

Turning to the out-of-sample results, the "Out-of-Sample Fcst." column of Sample 2 summarizes the results from using the coefficients estimated in the first sub-sample to form portfolios in the second sub-sample. The in- and out-of-sample portfolios are remarkably similar in terms of the distribution of the portfolio weights. More importantly, the return

statistics are also very similar. For instance, the average return and volatility are 24.5% and 16.3% when using the optimal coefficients (in-sample) versus 26.5% and 22.4% when using the coefficients estimated from Sample 1 (out-of-sample). The resulting certainty equivalents are 17.7% in-sample and 12.1% out-of-sample. It is astonishing that using portfolio policy coefficients estimated with a delay as long as 20 years results in less than a 6% reduction in certainty equivalent, compared to the (in practice unattainable) in-sample optimum. Moreover, the average characteristics of the in- and out-of-sample portfolios are similar, both with a tilt toward firms that are small, have high book-to-market ratios, and had large returns in the previous year. Notice that, although the sign of the average size is different for the two portfolios, both are considerably smaller than the average size of the market portfolio, and therefore both portfolios reflect a bias toward small firms.

The "Out-of-Sample Fcst." column of Sample 1 summarizes the results from using in the first sub-sample the coefficients estimated from the second sub-sample. The results are qualitatively the same as above. The in- and out-of-sample portfolios are remarkably similar, in terms of their weights, average characteristics, and return performance. The loss in certainty equivalent from not optimizing in-sample is only 3.8% in this case.

We conclude from these results that our approach is likely to perform as well out-of-sample as our in-sample analysis suggests. The same conclusion can be drawn from Panel B of Table 8, which reports the results for the portfolio subject to short-sale constraints. Since the constrained portfolios take less aggressive size, value, and momentum bets, the margin for error is smaller still. As a result, the certainty equivalent loss from not optimizing in-sample is a mere 1.2% for the first sub-sample and 1.7% for the second sub-sample.

4   Conclusion

We proposed a novel approach to optimizing large-scale equity portfolios. The portfolio weight in each stock is modeled as a function of the firm's characteristics, such as its market capitalization, book-to-market ratio, and lagged return. The coefficients of this function are found by optimizing the investor's average utility of the portfolio's return over a given sample period. We argued that our approach is computationally simple, easily modified and extended, produces sensible portfolio weights, and offers robust performance in- and out-of-sample.

We illustrated the many features of our approach through an empirical application to the universe of stocks in the CRSP-Compustat dataset. Our empirical results document the importance of the firm's market capitalization, book-to-market ratio, and one-year lagged return for explaining deviations of the optimal portfolio for a CRRA investor from the market. Relative to market capitalization weights, the optimal portfolio (with and without short-sale constraints) allocates considerably more wealth to stocks of small firms, firms with high book-to-market ratios (value firms), and firms with large positive lagged returns (past winners). With a relative risk aversion of five, the certainty equivalent gain from incorporating the firm characteristics, relative to holding the market portfolio, is an annualized 10% without short-sale constraints and 3% with short-sale constraints. Without short-sale constraints, the benefits are even greater when we allow for interactions between characteristics (certainty equivalent gain of 18%), especially interactions between lagged returns and book-to-market ratios, or when we allow for the coefficients of the portfolio policy to depend on the slope of the yield curve (certainty equivalent gain of 12%). With short-sale constraints, these extensions yield relatively little economic benefit compared to the base case. We used an out-of-sample experiment to show that these gains in expected utility are not the result of in-sample overfitting.

We compared the utility derived from the optimal portfolio with the utility obtained from holding the market. This comparison provides a test of equilibrium in the stock market with a representative investor who has the given utility function. In general, we rejected the hypothesis of equilibrium pricing, especially if the investor is allowed to take into account interactions between characteristics or to condition the portfolio policy on macro variables. However, if the representative investor is restricted to hold only long positions, we cannot reject the optimality of the market portfolio. Short-sale constraints therefore offer a justification for the size, value, and momentum anomalies. We also found that increasing the level of risk aversion eliminates the exposure of the optimal portfolio to size and momentum, suggesting that the premium of these stocks may indeed bear some relation to risk, but does not change the exposure to book to market. Finally, long-horizon investors still want to be exposed to value and momentum, contrary to claims that these premia represent compensation for intertemporal risks.

Our idea can easily be applied to other asset classes. For instance, we could use a similar approach to form bond portfolios based on the characteristics of the bond (e.g., duration, convexity, coupon rate, credit rating, leverage, etc) or to form currency portfolios

based on the characteristics of each country pair (e.g., interest rate and inflation differentials, trade balance, etc).

References

Ait-Sahalia, Yacine, and Michael W. Brandt, 2001, Variable selection for portfolio choice, *Journal of Finance* 56, 1297–1351.

———, 2002, Portfolio and consumption choice with option-implied state prices, Working Paper, Princeton University.

Alexander, Gordon, and Alexandre Baptista, 2002, Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis, *Journal of Economic Dynamics and Control* 26, 1159–1193.

Ang, Andrew, Geert Bekaert, and Jun Liu, 2003, Why stocks may disappoint, *Journal of Financial Economics* forthcoming.

Bansal, Ravi, Robert Dittmar, and Christian Lundblad, 2004, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance* forthcoming.

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.

Basak, Suleyman, and Alexander Shapiro, 2001, Value-at-risk based risk management: optimal policies and asset prices, *Review of Financial Studies* 14, 371–405.

Benartzi, Shlomo, and Richael H. Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics* 110, 73–92.

Black, Fisher, and Robert Litterman, 1992, Global portfolio optimization, *Financial Analysts Journal* 48, 28–43.

Brandt, Michael W., 1999, Estimating portfolio and consumption choice: A conditional method of moments approach, *Journal of Finance* 54, 1609–1646.

———, 2004, Portfolio choice problems, in Yacine Ait-Sahalia, and Lars P. Hansen, ed.: *Handbook of Financial Econometrics.*

———, and Pedro Santa-Clara, 2004, Dynamic portfolio selection by augmenting the asset space, Working Paper, UCLA.

Brennan, Michael, and Yihong Xia, 2004, Risk and valuation under an intertemporal asset pricing model, *Journal of Business* forthcoming.

Browne, Sid, 1999, Beating a moving target: Optimal portfolio strategies for outperforming a stochastic benchmark, *Finance and Stochastics* 3, 275–294.

Campbell, John, and Tuomo Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review* forthcoming.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

———, and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.

Chan, K.C., Jason Karceski, and Josef Lakonishok, 1998, The risk and return from factors, *Journal of Financial and Quantitative Analysis* 33, 159–188.

———, 1999, On portfolio optimization: Forecasting covariances and choosing the risk model, *Review of Financial Studies* 12, 937–974.

Cooper, Michael, Huseyin Gulen, and Maria Vassalou, 2000, Investing in size and book-to-market portfolios using information about the macroeconomy: Some new trading rules, Working Paper, Purdue University.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Daniel, Kent D., and Sheridan Titman, 1999, Market efficiency in an irrational world, *Financial Analyst Journal* 55, 24–50.

Fama, Eugene F., 1990, Stock returns, expected returns, and real activity, *Journal of Finance* 45, 1089–1108.

———, and Kenneth R. French, 1988, Permanent and temporary components of stock prices, *Journal of Political Economy* 96, 246–273.

———, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

———, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

———, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Frost, Peter A., and James E. Savarino, 1986, An empirical Bayes approach to efficient portfolio selection, *Journal of Financial and Quantitative Analysis* 21, 293–305.

———, 1988, For better performance: Constrain portfolio weights, *Journal of Portfolio Management* 15, 29–34.

Gomes, Francisco, 2003, Portfolio choice and trading volume with loss-averse investors, *Journal of Business* forthcoming.

Grossman, Sanford J., and Jean-Luc Vila, 1989, Portfolio insurance in complete markets: A note, *Journal of Business* 62, 473–476.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1053.

Heaton, John, and Deborah Lucas, 2000, Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance* 55, 1163–1198.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 257–286.

Jagannathan, Ravi, and Tongshu Ma, 2002, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1684.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, *Journal of Finance* 51, 3–53.

Jobson, J. David, and Bob Korkie, 1980, Estimation of Markowitz efficient portfolios, *Journal of the American Statistical Association* 75, 544–554.

———, 1981, Putting markowitz theory to work, *Journal of Portfolio Management* 7, 70–74.

Jorion, Philippe, 1986, Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21, 279–292.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Ledoit, Olivier, and Michael Wolf, 2003a, Honey, I shrunk the sample covariance matrix, *Journal of Portfolio Management* forthcoming.

———, 2003b, Improved estimation of the covariance matrix of returns with an application to portfolio selection, *Journal of Empirical Finance* forthcoming.

Markowitz, Harry M., 1952, Portfolio selection, *Journal of Finance* 7, 77–91.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Michaud, Robert O., 1989, The markowitz optimization enigma: Is optimized optimal?, *Financial Analyst Journal* 45, 31–42.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Nigmatullin, Eldar A., 2003, Bayesian model averaging for moment conditions models, Working Paper, University of Wisconsin-Madison.

Pastor, Lubos, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 50, 179–223.

———, and Robert F. Stambaugh, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

———, 2002, Investing in equity mutual funds, *Journal of Financial Economics* 63, 351–380.

Politis, Dimitris N., and Joseph P. Romano, 1994, The stationary bootstrap, *Journal of the American Statistical Association* 89, 1303–1313.

Roy, A.D., 1952, Safety first and the holding of assets, *Econometrica* 20, 431–439.

Stambaugh, Robert, 1997, Analyzing investments whose histories differ in length, *Journal of Financial Economics* 54, 375–421.

Tepla, Lucie, 2001, Optimal investment with minimum performance constraints, *Journal of Economic Dynamics and Control* 25, 1629–1645.

Valkanov, Rossen, 2003, Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics* 68, 201–232.

Appendix: Data

For each firm in the CRSP-Compustat dataset, we construct several variables at the end of fiscal years 1964 to 2002. The first full year of data, 1963, is used to construct lagged values. The exact fiscal year end dates are from CRSP. We use the following quantities in the definition of the variables (Compustat data item numbers are in parenthesis): total assets (6); liabilities (181); preferred stock value (10, 56, or 130, in that order, or, otherwise, zero); balance sheet deferred taxes and investment tax credits (35, otherwise zero); price per share (199, otherwise taken from CRSP); and shares outstanding (25, otherwise taken from CRSP). If total assets, liabilities, price, and shares outstanding are missing, the observation is not included in the dataset. Then, we define book equity (BE) as equal to total assets minus liabilities plus balance sheet deferred taxes and investment tax credits minus preferred stock value; market equity (ME) as equal to price per share times shares outstanding; book-to-market (btm) as equal to the log of one plus book equity divided by market equity. We omit firms with negative book-to-market ratio. Log market equity (me) is computed as the log of market equity.

The monthly firm returns are obtained from CRSP. We allow a minimum of six month lag between the fiscal year end of the above accounting variables and the returns to ensure that the information from the firms' annual reports would have been publicly available at the time of portfolio formation. From CRSP, we also compute the trailing twelve-month return (mom), defined as the monthly compounded return between months $t - 13$ and $t - 2$. After all variables have been created, we eliminate the smallest 20 percent of firms (i.e., firms in the lowest 20 percentile of me).

Table 1: Base Case

This table shows estimates of the portfolio policy with three characteristics: size (me), book-to-market ratio (btm), and momentum (mom), specified in equation (3). In panel "All Stocks," we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. Panel "Top 500 Stocks" includes only the largest 500 companies in each month. The columns labeled "Val.Weighted" and "Opt." display statistics of the market-capitalization weighted portfolio and the optimal portfolio, respectively. The first set of rows shows the estimated coefficients of the portfolio policy with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second set of rows shows statistics of the portfolio weights, averaged across time. These statistics include the average absolute portfolio weight ($|w_i|$), the average minimum and maximum portfolio weights (max w_i and min w_i), the average sum of negative weights in the portfolio ($\sum w_i I(w_i < 0)$), and the average fraction of negative weights in the portfolio ($\sum I(w_i \leq 0)/N$). The third set of rows displays average portfolio return statistics: average return, standard deviation, and Sharpe ratio of returns ($\bar{r}$, $\sigma(r)$, and SR), the certainty equivalent return, the alpha, beta, and volatility of idiosyncratic shocks of a market model regression, and the information ratio. The final set of rows displays the average normalized characteristics of the portfolio. The average risk-free rate in the sample is 0.061 (annualized).

Variable	All Stocks		Top 500 Stocks	
	Val.Weighted	Opt.	Val.Weighted	Opt.
θ_{me}	—	-1.220	—	-1.389
	—	(0.547)	—	(1.145)
θ_{btm}	—	3.466	—	1.557
	—	(0.922)	—	(0.725)
θ_{mom}	—	2.000	—	1.709
	—	(0.742)	—	(0.561)
Wald p-value	—	0.000	—	0.004
$|w_i| \times 100$	0.035	0.121	0.200	0.407
max $w_i \times 100$	4.625	4.363	5.343	4.226
min $w_i \times 100$	0.000	-0.323	0.037	-1.126
$\sum w_i I(w_i < 0)$	0.000	-1.238	0.000	-0.518
$\sum I(w_i \leq 0)/N$	0.000	0.468	0.000	0.358
$\bar{r}$	0.120	0.244	0.117	0.200
$\sigma(r)$	0.161	0.190	0.158	0.176
SR	0.362	0.964	0.349	0.790
CE(r)	0.052	0.153	0.052	0.115
α	—	0.157	—	0.100
β	—	0.440	—	0.704
$\sigma(\epsilon)$	—	0.177	—	0.136
IR	—	0.890	—	0.734
me	2.095	-0.118	1.527	-0.136
btm	-0.451	3.296	-0.240	1.430
mom	0.012	1.839	-0.026	1.622

Table 2: Short-Sale Constraint

This table shows estimates of the portfolio policy with three characteristics: size (me), book-to-market ratio (btm), and momentum (mom). The portfolio weights, specified in equation (14), are restricted to be non-negative (long-only positions). In panel "All Stocks," we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. Panel "Top 500 Stocks" includes only the largest 500 companies in each month. The columns labeled "Val.Weighted" and "Opt." display statistics of the market-capitalization weighted portfolio and the optimal portfolio, respectively. The first set of rows shows the estimated coefficients of the portfolio policy with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

Variable	All Stocks Val.Weighted	Pos.Opt.	Top 500 Stocks Val.Weighted	Pos.Opt.
θ_{me}	—	-0.182	—	-0.871
	—	(1.037)	—	(1.045)
θ_{btm}	—	2.043	—	1.445
	—	(1.113)	—	(1.114)
θ_{mom}	—	2.358	—	2.359
	—	(1.199)	—	(1.411)
Wald p-value	—	0.114	—	0.109
$\lvert w_i \rvert \times 100$	0.035	0.035	0.200	0.200
max $w_i \times 100$	4.625	2.455	5.343	2.988
min $w_i \times 100$	0.000	0.000	0.037	0.000
$\sum w_i I(w_i < 0)$	0.000	0.000	0.000	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.439	0.000	0.375
$\bar{r}$	0.120	0.173	0.117	0.169
$\sigma(r)$	0.161	0.180	0.158	0.170
SR	0.362	0.620	0.349	0.632
CE(r)	0.052	0.085	0.052	0.091
α	—	0.054	—	0.054
β	—	0.977	—	0.971
$\sigma(\epsilon)$	—	0.088	—	0.073
IR	—	0.620	—	0.740
me	2.095	0.632	1.527	0.228
btm	-0.451	0.451	-0.240	0.433
mom	0.012	0.886	-0.026	0.998

Table 3: Interactions

This table shows estimates of the portfolio policy with six characteristics: size (me), book-to-market ratio (btm), momentum (mom), and the cross product of the characteristics me × btm, me × mom, and btm × mom. In the estimation, we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. The column labeled "Val.Weighted" displays the statistics of the market-capitalization weighted portfolio, as a benchmark. The next two columns, "Opt." and "Pos.Opt." characterize the optimal and positive optimal (long-only) portfolios, respectively. The first set of rows shows the estimated coefficients of the portfolio policies (3) and (14) with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

| | All Stocks | | |
Variable	Val.Weighted	Opt.	Pos.Opt.
θ_{me}	—	-0.744	-0.836
	—	(0.680)	(1.377)
θ_{btm}	—	4.444	5.596
	—	(1.062)	(1.194)
θ_{mom}	—	4.631	4.565
	—	(1.565)	(1.353)
$\theta_{\text{me}\times\text{btm}}$	—	-1.727	3.325
	—	(0.894)	(1.178)
$\theta_{\text{me}\times\text{mom}}$	—	4.718	5.346
	—	(1.679)	(1.849)
$\theta_{\text{btm}\times\text{mom}}$	—	4.249	4.000
	—	(1.837)	(1.384)
Wald p-value	—	0.000	0.000
$|w_i| \times 100$	0.035	0.186	0.035
max $w_i \times 100$	4.625	4.926	1.684
min $w_i \times 100$	0.000	-2.420	0.000
$\sum w_i I(w_i < 0)$	0.000	-1.939	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.503	0.511
$\bar{r}$	0.120	0.428	0.200
$\sigma(r)$	0.161	0.278	0.186
SR	0.362	1.318	0.746
CE(r)	0.052	0.231	0.106
α	—	0.324	0.083
β	—	0.742	0.962
$\sigma(\epsilon)$	—	0.252	0.104
IR	—	1.283	0.795
me	2.095	-0.003	0.369
btm	-0.451	4.714	0.684
mom	0.012	1.446	0.620
me × btm	-1.142	-5.398	-0.254
me × mom	0.017	1.864	0.181
btm × mom	-0.014	3.281	0.718

Table 4: Conditioning on the Slope of the Yield Curve

This table shows estimates of the portfolio policy with the product of three characteristics, size (me), book-to-market ratio (btm), and momentum (mom), and an indicator function of the sign of the slope of the yield curve, $I(\text{tsp} \geq 0)$ or $I(\text{tsp} < 0)$. In the estimation, we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. The column labeled "Val.Weighted" displays the statistics of the market-capitalization weighted portfolio, as a benchmark. The next two columns, "Opt." and "Pos.Opt." characterize the optimal and positive optimal (long-only) portfolios, respectively. The first set of rows shows the estimated coefficients of the portfolio policies with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

| | | All Stocks | |
Variable	Val.Weighted	Opt.	Pos.Opt.
$\theta_{\text{me} \times I(\text{tsp}>0)}$	—	-1.918	-1.316
	—	(0.707)	(1.120)
$\theta_{\text{me} \times I(\text{tsp}\leq 0)}$	—	0.952	2.879
	—	(1.197)	(2.044)
$\theta_{\text{btm} \times I(\text{tsp}>0)}$	—	3.270	2.292
	—	(1.101)	(1.124)
$\theta_{\text{btm} \times I(\text{tsp}\leq 0)}$	—	4.965	3.773
	—	(2.064)	(1.724)
$\theta_{\text{mom} \times I(\text{tsp}>0)}$	—	2.128	3.101
	—	(0.911)	(1.710)
$\theta_{\text{mom} \times I(\text{tsp}\leq 0)}$	—	3.499	2.343
	—	(1.613)	(1.874)
Wald p-value	—	0.000	0.007
$|w_i| \times 100$	0.035	0.136	0.035
max $w_i \times 100$	4.625	4.370	2.033
min $w_i \times 100$	0.000	-0.389	0.000
$\sum w_i I(w_i < 0)$	0.000	-1.411	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.475	0.464
$\bar{r}$	0.120	0.273	0.181
$\sigma(r)$	0.161	0.203	0.182
SR	0.362	1.042	0.658
CE(r)	0.052	0.171	0.093
α	—	0.191	0.064
β	—	0.357	0.947
$\sigma(\epsilon)$	—	0.196	0.100
IR	—	0.977	0.644
me $\times$ I(tsp>0)	1.616	-0.556	0.120
me $\times$ I(tsp $\leq$ 0)	0.479	0.340	0.242
btm $\times$ I(tsp>0)	-0.346	2.565	0.459
btm $\times$ I(tsp $\leq$ 0)	-0.105	0.909	0.095
mom $\times$ I(tsp>0)	0.027	1.534	0.818
mom $\times$ I(tsp $\leq$ 0)	-0.015	0.693	0.098

Table 5: Risk Aversion

This table shows estimates of the portfolio policy with three characteristics, size (me), book-to-market ratio (btm), and momentum (mom) using different utility functions. In additions to the $\gamma = 5$ case, reported as a benchmark, we display the results for $\gamma = 1$, $\gamma = 100$, as well as the value-weighted case, as a reference. In the estimation, we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. The columns labeled "Optimal" and "Positive Optimal" characterize the optimal and positive optimal (long-only) portfolios, respectively. The first set of rows shows the estimated coefficients of the portfolio policies with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1. The certainty equivalent returns of the market portfolio for $\gamma=1$, $\gamma=5$ and $\gamma = 100$ are 0.107, 0.052 and -1.976 (annualized), respectively.

	Val.Weighted	Optimal			Positive Optimal		
		$\gamma=1$	$\gamma=5$	$\gamma=100$	$\gamma=1$	$\gamma=5$	$\gamma=100$
θ_{me}	—	-6.109	-1.220	0.146	-3.539	-0.182	-0.032
	—	(2.881)	(0.547)	(0.227)	(2.582)	(1.037)	(0.128)
θ_{btm}	—	6.970	3.466	5.025	0.132	2.043	0.114
	—	(3.548)	(0.922)	(0.946)	(3.097)	(1.113)	(0.162)
θ_{mom}	—	7.380	2.000	0.483	13.368	2.358	-0.048
	—	(2.915)	(0.742)	(0.195)	(3.604)	(1.199)	(0.087)
Wald p-value	—	0.000	0.000	0.000	0.000	0.114	0.067
$|w_i| \times 100$	0.035	0.320	0.121	0.146	0.035	0.035	0.035
max $w_i \times 100$	4.625	4.433	4.363	4.497	1.397	2.455	4.574
min $w_i \times 100$	0.000	-0.960	-0.323	-0.350	0.000	0.000	0.000
$\sum w_i I(w_i < 0)$	0.000	-4.026	-1.238	-1.594	0.000	0.000	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.520	0.468	0.477	0.564	0.439	0.136
$\bar{r}$	0.120	0.546	0.244	0.190	0.210	0.173	0.121
$\sigma(r)$	0.161	0.598	0.190	0.187	0.241	0.180	0.159
SR	0.362	0.811	0.964	0.690	0.617	0.620	0.375
CE(r)	—	0.363	0.153	-0.976	0.181	0.085	-1.965
α	—	0.507	0.157	0.109	0.084	0.054	0.002
β	—	-0.375	0.440	0.345	1.108	0.977	0.989
$\sigma(\epsilon)$	—	0.595	0.177	0.179	0.163	0.088	0.007
IR	—	0.851	0.890	0.609	0.517	0.620	0.309
me	2.095	-5.932	-0.118	0.744	-0.178	0.632	2.001
btm	-0.451	8.043	3.296	4.508	-0.086	0.451	-0.333
mom	0.012	6.958	1.839	0.290	2.082	0.886	-0.020

Table 6: Minimum Variance and Maxiumum Sharpe Ratio

This table shows estimates of the portfolio policy with three characteristics, size (me), book-to-market ratio (btm), and momentum (mom) used to minimize the variance ("Var") and to maximize the Sharpe ratio ("SR"). In the estimation, we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. The columns labeled "Optimal" and "Positive Optimal" characterize the optimal and positive optimal (long-only) portfolios, respectively. The first set of rows shows the estimated coefficients of the portfolio policies with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

Variable	Val.Weighted	Optimal		Positive Optimal	
		Var	SR	Var	SR
θ_{me}	—	0.006	-1.117	-0.089	0.883
	—	(0.196)	(0.587)	(0.152)	(5.544)
θ_{btm}	—	1.887	3.303	0.461	7.462
	—	(0.311)	(1.092)	(0.131)	(3.535)
θ_{mom}	—	0.455	2.415	0.111	11.209
	—	(0.297)	(1.037)	(0.138)	(5.557)
Wald p-value	—	0.000	0.004	0.001	0.057
$\|w_i\| \times 100$	0.035	0.070	0.121	0.035	0.035
max $w_i \times 100$	4.625	4.568	4.382	4.280	1.134
min $w_i \times 100$	0.000	-0.142	-0.336	0.000	0.000
$\sum w_i I(w_i < 0)$	0.000	-0.523	-1.242	0.000	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.418	0.466	0.284	0.501
$\bar{r}$	0.120	0.155	0.251	0.131	0.198
$\sigma(r)$	0.161	0.137	0.195	0.156	0.209
SR	0.362	0.680	0.973	0.444	0.651
CE(r)	—	—	—	—	—
α	—	0.050	0.164	0.013	0.077
β	—	0.742	0.447	0.963	1.021
$\sigma(\epsilon)$	—	0.069	0.182	0.023	0.130
IR	—	0.731	0.900	0.584	0.589
me	2.095	1.545	0.043	1.705	0.021
btm	-0.451	1.414	3.085	-0.053	0.617
mom	0.012	0.388	2.263	0.063	1.484

Table 7: Long-Horizon Portfolio Policy

This table shows estimates of a long-horizon portfolio policy with three characteristics, size (me), book-to-market ratio (btm), and momentum (mom) over horizons of K periods, where K is equal to 1, 12, and 60 months. For each horizon, the estimates are obtained by maximizing average utility over the cumulative, non-overlapping K-period portfolio returns. The results for the "Val.Weighted" and the $K = 1$ case are identical to those presented in the above tables and are shown here for reference. Panels A and B display the results from the unrestricted and positive (long-only) weights. In the estimation, we use all stocks in the merged CRSP-Compustat database from January 1964 through December 2002. The first set of rows shows the estimated coefficients of the portfolio policy with bootstrapped standard errors in parentheses. The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

| | Panel A: Unrestricted Optimal Portfolio | | | |
	Val.Weighted	K=1	K=12	K=60
θ_{me}	—	-1.220	-0.843	-1.494
	—	(0.547)	(0.841)	(1.523)
θ_{btm}	—	3.466	3.281	3.857
	—	(0.922)	(1.221)	(1.847)
θ_{mom}	—	2.000	2.472	4.371
	—	(0.742)	(1.338)	(1.478)
Wald p-value	—	0.000	0.010	0.015
$|w_i| \times 100$	0.035	0.121	0.120	0.135
max $w_i \times 100$	4.625	4.363	4.417	3.959
min $w_i \times 100$	0.000	-0.323	-0.337	-0.414
$\sum w_i I(w_i < 0)$	0.000	-1.238	-1.229	-1.800
$\sum I(w_i \leq 0)/N$	0.000	0.468	0.463	0.482
$\bar{r}$	0.120	0.244	0.246	0.312
$\sigma(r)$	0.161	0.190	0.191	0.280
SR	0.362	0.964	0.969	0.896
CE(r)	0.052	0.153	0.160	0.329
α	—	0.157	0.158	0.234
β	—	0.440	0.457	0.241
$\sigma(\epsilon)$	—	0.177	0.177	0.277
IR	—	0.890	0.895	0.845
me	2.095	-0.118	0.325	-0.370
btm	-0.451	3.296	2.979	3.694
mom	0.012	1.839	2.327	4.202

| | Panel B: Long-Only Optimal Portfolio | | | |
	Val.Weighted	K=1	K=12	K=60
θ_{me}	—	-0.182	2.924	1.168
	—	(1.037)	(1.866)	(0.669)
θ_{btm}	—	2.043	4.343	1.455
	—	(1.113)	(3.913)	(1.147)
θ_{mom}	—	2.358	5.106	3.168
	—	(1.199)	(2.290)	(1.105)
Wald p-value	—	0.114	0.121	0.130
$|w_i| \times 100$	0.035	0.035	0.035	0.029
max $w_i \times 100$	4.625	2.455	1.495	1.998
min $w_i \times 100$	0.000	0.000	0.000	0.000
$\sum w_i I(w_i < 0)$	0.000	0.000	0.000	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.439	0.471	0.467
$\bar{r}$	0.120	0.173	0.174	0.162
$\sigma(r)$	0.161	0.180	0.181	0.184
SR	0.362	0.620	0.624	0.549
CE(r)	0.052	0.085	0.096	0.134
α	—	0.054	0.055	0.041
β	—	0.977	0.997	1.012
$\sigma(\epsilon)$	—	0.085	0.095	0.074
IR	—	0.620	0.645	0.550
me	2.095	0.632	0.749	1.055
btm	-0.451	0.451	0.385	0.010
mom	0.012	0.886	1.000	1.053

Table continued from previous page.

Table 8: Out-of-Sample Performance

This table shows the out-of-sample performance of the portfolio policy with three characteristics, size (me), book to market ratio (btm), and momentum (mom). Panels A and B display the results from the unrestricted and positive (long-only) weights. The original data of all stocks in the merged CRSP-Compustat database is divided into two equal sub-samples, Sample 1 (January 1964 through June 1983) and Sample 2 (July 1983 through December 2002). In each sub-sample, we display in-sample and out-of-sample statistics. For the in-sample results, the column labeled "Val.Weighted" displays the statistics of the market-capitalization weighted portfolio, whereas the column "Opt." characterizes the optimal portfolio. The out-of-sample forecasts in Sample 1 are obtained using the in-sample estimates of the optimal investment policy function from Sample 2 in forming the optimal portfolio. Similarly, the out-of-sample forecasts in Sample 2 are obtained from the in-sample estimates in Sample 1. The optimal portfolio parameters are not re-estimated in the forecasts. The first set of rows shows the estimated coefficients of the portfolio policies with bootstrapped standard errors in parentheses. (Cont'd on next page.)

	Panel A: Unrestricted Optimal Portfolio					
	Sample 1			Sample 2		
	In-Sample		Out-of-Sample	In-Sample		Out-of-Sample
	Val.Weighted	Opt.	Fcst.	Val.Weighted	Opt.	Fcst.
θ_{me}	—	-1.707	-0.702	—	-0.702	-1.707
	—	(0.891)	—	—	(0.822)	—
θ_{btm}	—	2.891	4.166	—	4.166	2.891
	—	(1.442)	—	—	(1.194)	—
θ_{mom}	—	3.111	1.346	—	1.346	3.111
	—	(1.149)	—	—	(0.931)	—
Wald p-value	—	0.003	—	—	0.001	—
$|w_i| \times 100$	0.049	0.176	0.180	0.021	0.079	0.075
max $w_i \times 100$	6.119	5.661	5.802	3.130	3.019	2.990
min $w_i \times 100$	0.000	-0.516	-0.472	0.000	-0.181	-0.205
$\sum w_i I(w_i < 0)$	0.000	-1.307	-1.359	0.000	-1.363	-1.289
$\sum I(w_i \leq 0)/N$	0.000	0.470	0.464	0.000	0.480	0.478
$\bar{r}$	0.103	0.293	0.201	0.136	0.245	0.265
$\sigma(r)$	0.155	0.233	0.194	0.166	0.163	0.224
SR	0.239	0.965	0.689	0.417	1.160	0.936
CE(r)	0.043	0.151	0.113	0.061	0.177	0.121
α	—	0.197	0.141	—	0.177	0.197
β	—	0.795	0.709	—	0.154	0.153
$\sigma(\epsilon)$	—	0.198	0.194	—	0.162	0.223
IR	—	0.994	0.728	—	1.095	0.884
me	2.078	-0.586	0.026	2.111	0.322	-0.228
btm	-0.528	2.732	3.784	-0.375	3.952	2.913
mom	-0.070	2.890	1.044	0.094	1.293	3.034

Table continued on next page ...

(Cont'd from previous page.) The bootstrapped p-value of the Wald test under the null hypothesis that the parameter estimates are jointly equal to zero is also displayed. For the out-of-sample results, we display the coefficients that were used in producing the forecasts. The second, third, and fourth set of rows show the statistic of the portfolio weights, returns, and normalized characteristics, respectively, which are described in detail in Table 1.

	Panel B: Long-Only Optimal Portfolio					
	Sample 1			Sample 2		
	In-Sample		Out-of-Sample	In-Sample		Out-of-Sample
	Val.Weighted	Opt.	Fcst.	Val.Weighted	Opt.	Fcst.
θ_{me}	—	-0.746	0.036	—	0.036	-0.746
	—	(1.875)	—	—	(1.191)	—
θ_{btm}	—	3.198	1.887	—	1.887	3.198
	—	(2.029)	—	—	(1.283)	—
θ_{mom}	—	4.767	1.243	—	1.243	4.767
	—	(2.500)	—	—	(1.119)	—
Wald p-value	—	0.073	—	—	0.492	—
$\lvert w_i \rvert \times 100$	0.049	0.049	0.049	0.021	0.021	0.021
max $w_i \times 100$	6.119	2.186	3.804	3.130	1.882	1.185
min $w_i \times 100$	0.000	0.000	0.000	0.000	0.000	0.000
$\sum w_i I(w_i < 0)$	0.000	0.000	0.000	0.000	0.000	0.000
$\sum I(w_i \leq 0)/N$	0.000	0.479	0.408	0.000	0.431	0.479
$\bar{r}$	0.103	0.215	0.167	0.136	0.150	0.164
$\sigma(r)$	0.155	0.211	0.178	0.166	0.157	0.188
SR	0.239	0.701	0.557	0.417	0.606	0.582
CE(r)	0.043	0.099	0.087	0.061	0.080	0.063
α	—	0.107	0.104	—	0.025	0.039
β	—	1.156	1.074	—	0.862	0.869
$\sigma(\epsilon)$	—	0.111	0.179	—	0.065	0.121
IR	—	0.966	0.583	—	0.388	0.322
me	2.078	0.099	0.848	2.111	0.913	0.200
btm	-0.528	0.503	0.442	-0.375	0.541	0.613
mom	-0.070	1.220	0.437	0.094	0.541	1.313

Table continued from previous page.

Figure 1: Summary Statistics of Characteristics

The figure displays cross-sectional means and standard deviations of the firm characteristics *me*, *btm*, and *mom* in every month from January 1964 to December 2002. For each month and firm, the characteristics are *me*, defined as the log of market equity, *btm*, defined as the log of one plus the ratio of book equity divided by market equity, and *mom*, defined as the lagged 12 month return. The reported means and standard deviations are computed across firms at each point in time.
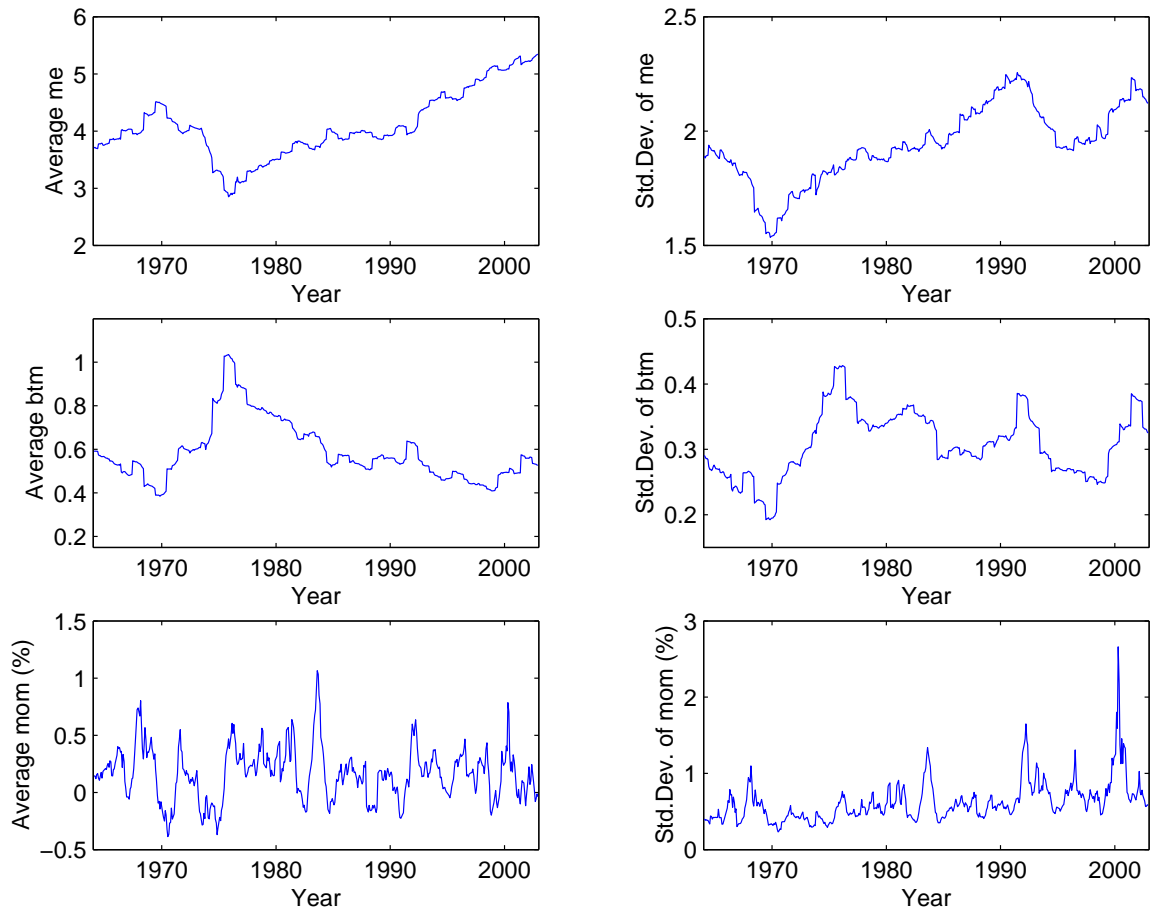
Figure 2: Portfolio Characteristics over Time

The figure displays the portfolio characteristics of policy function (3) using size, book-to-market and momentum as firm-specific characteristics. The estimates of θ are obtained using all available CRSP-Compustat stocks from January 1964 to December 2002. The utility function is specified with $\gamma = 5$. The average values of these characteristics are reported in the last three lines of Table 1.