University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 238

Efficient Weighting: A More Powerful Test for Cross-Sectional Anomalies

Olivier Ledoit, Michael Wolf and Zhao Zhao

First version: December 2016
This version: February 2018

Efficient Weighting:
A More Powerful Test for Cross-Sectional Anomalies*

Olivier Ledoit
Department of Economics
University of Zurich
CH-8032 Zurich, Switzerland
olivier.ledoit@econ.uzh.ch

Michael Wolf
Department of Economics
University of Zurich
CH-8032 Zurich, Switzerland
michael.wolf@econ.uzh.ch

Zhao Zhao
School of Economics
Huazhong University of Science and Technology
Wuhan, Hubei, China
zhaozhao@hust.edu.cn

First version: December 2016
This version: February 2018

Abstract

Many researchers seek factors that predict the cross-section of stock returns. The standard methodology sorts stocks according to their factor scores into quantiles and forms a corresponding long-short portfolio. Such a course of action ignores any information on the covariance matrix of stock returns. Historically, it has been difficult to estimate the covariance matrix for a large universe of stocks. We demonstrate that using the recent DCC-NL estimator of Engle et al. (2017) substantially enhances the power of tests for cross-sectional anomalies: On average, 'Student' t-statistics more than double.

KEY WORDS: Cross-section of returns, dynamic conditional correlations, GARCH, Markowitz portfolio selection, nonlinear shrinkage.

JEL CLASSIFICATION NOS: C13, C58, G11.

*We thank Rob Engle for helpful comments.

1 Introduction

The search for factors that predict the cross-section of stock returns generates an abundant literature. Instead of "factors", some authors may use alternative terms such as signals, predictors, characteristics, anomalies, cross-sectional patterns, forecasting variables, etc. What we mean specifically is a function of historical data that can explain the cross-section of subsequent stock returns: discriminate between the stocks that will tend to outperform their peers and the ones that will tend to underperform their peers. Both Green et al. (2013) and Harvey et al. (2016) find more than 300 articles and factors in this strand of literature.

At least since Fama and French (1992), the preferred method for establishing the validity of factors has been to construct portfolios based on sorting. For example, one can form a dollar-neutral long-short portfolio by going long the stocks that are in the top quintile according to their factor scores, and short the stocks in the bottom quintile. Instead of quintiles, some authors may prefer terciles, deciles, etc. The portfolio is then held for a certain period of time, at which point it is rebalanced according to freshly updated factor data. This procedure generates a time series of portfolio returns. The factor is deemed successful if the average portfolio return exceeds some benchmark, generally zero percent, at a suitable level of statistical significance. Thus, the central quantity is the 'Student' t-statistic of the long-short portfolio return. This test is called *predictive* in the sense that, at any point in time, portfolio construction rules involve only data that was acquired earlier. Such investment strategies are realistic and can be implemented by a quantitative fund manager.

This *status quo* poses a conundrum: How come we have a quantitative investment strategy that does not employ the covariance matrix of asset returns? Indeed, the historical foundation of finance as a mathematically rigorous discipline can be traced back to the discovery of Markowitz (1952) portfolio selection. He proved that optimal portfolio weights depend not only on (a factor that proxies for) the first moment of returns, but also on the second moment: the covariance matrix — or, to be precise, its inverse. A more powerful test for cross-sectional anomalies can be designed by replacing the traditional portfolio construction rule based on sorting with a more efficient one that incorporates the (inverse) covariance matrix, at least in theory.

From theory to practice there is a gap: The true covariance matrix is unobservable; therefore, it needs to be estimated somewhow. At the epoch when the standard procedure for testing factors crystallized around sorting, there was no covariance matrix estimator that could cope with inversion in large dimensions. Indeed, Michaud (1989) described portfolio optimization as an "error-maximization procedure". Ledoit and Wolf (2004b) showed that the textbook estimator, the sample covariance matrix, is ill-conditioned when the dimension is not negligible with respect to sample size: inverting it amplifies any estimation error. This unfortunate behavior is pushed to a numerical extreme when the number of stocks exceeds the number of time series observations, at which point the supposedly optimal portfolio weights blow up to plus or minus infinity for no reason whatsoever — which violates economic sense. Even with two years of daily data at hand, this systemic breakdown happens as soon as we consider the universe of S&P 500 constituents.

Abandoning mean-variance optimization for portfolio selection would amount to 'throwing the baby out with the bathwater'. The way forward instead is to consider an improved covariance matrix estimator that fixes the weaknesses of the sample covariance matrix, so that the profession as a whole can upgrade from sorting-based portfolios to a more efficient weighting scheme. This is the purpose of the present paper. As it turns out, covariance matrix estimation has been an active field of research over the recent years. Substantive progress has been achieved in two complementary directions.

The first direction is time series. Variances and covariances move over time, and they need to be tracked accordingly, which the sample covariance matrix is not geared to do. Early success in this area was achieved in the univariate case by the ARCH model of Engle (1982), followed by generalizations such as the GARCH model of Bollerslev (1986), and too many others to review here. Extension to the multivariate case, however, has been slowed down by the curse of dimensionality. The main breakthroughs in this challenging area have been: (i) volatility targeting (Engle and Mezrich, 1996); (ii) the Dynamic Conditional Correlation (DCC) model of Engle (2002); and (iii) composite likelihood estimation (Pakel et al., 2017). Together they solve the difficulties attributable to the time-varying aspects of the covariance matrix — but only provided that cross-sectional issues intrinsic to the estimation of large-dimensional unconditional covariance matrices can be fixed on their own terms.

This leads us to the second direction where substantive progress has been accomplished: the cross-section. Stein (1986) showed that, absent *a priori* structural information, the eigenvectors of the sample covariance matrix can be preserved, but its eigenvalues must be nonlinearly shrunk towards their cross-sectional average due to systematic in-sample overfitting. He also hinted that a nonstandard asymptotic theory might shed some light: large-dimensional asymptotics, where the matrix dimension is assumed to go to infinity along with the sample size. However, much work remained to be done by a variety of authors such as Silverstein and Bai (1995) until Ledoit and Péché (2011) derived the theoretically optimal nonlinear shrinkage formula, and Ledoit and Wolf (2012, 2015) developed a statistical implementation that works even when dimension exceeds sample size: the NonLinear (NL) shrinkage estimator of the unconditional covariance matrix.

The state-of-the-art developments in these two streams of covariance matrix estimation literature are brought together for the first time in the DCC-NL model of Engle et al. (2017). These authors examine the performance of mean-variance efficient portfolios subject to two linear constraints: the unit vector (for the global minimum variance portfolio) and the momentum factor. They find that, indeed, the DCC-NL estimator generates economically and statistically significant improvements in both cases.

There are two important differences between the present paper and Engle et al. (2017). First, we do not just look at two linear constraints in the mean-variance optimization problem but instead at a large ensemble of 60-plus different factors culled from the literature on cross-sectional anomalies. Second, we use dollar-neutral portfolios, whose weights sum up to zero, instead of fully-invested portfolios, whose weights sum up to one.

Our main original contribution is to demonstrate that using the DCC-NL estimator of the covariance matrix in a large investment universe multiplies the 'Student' t-statistics for cross-sectional anomaly detection, on average, by a factor of more than *two* relative to the *status quo*. Therefore, it is in everybody's interest to upgrade the theoretically and empirically underpowered portfolio-construction procedure based on sorting into quantiles.

The power boost from using DCC-NL is significant because it enables factor candidates that have a short history to get a chance at getting detected. Multiplying the t-statistic by two is equivalent to multiplying the number of years in the dataset by approximately four. Thus, if a given factor requires 40 years of historical data to achieve statistical significance with sorting, with DCC-NL the same factor can attain the same level of statistical significance in only ten years. This is especially relevant for all factors that are extracted from traffic on social networks, as these have only been active on a massive scale for a relatively small number of years. Given the explosion in data collection driven by the precipitous fall in storage cost per petabyte in recent years, this is just the tip of the iceberg: Big data is young data.

On a separate but equally important note, given that Harvey et al. (2016) claim that the significance threshold for t-statistics should be raised from two to three due to multiple-testing issues, it will be much harder for subsequent authors to meet this hurdle. Any candidate needs all the power boost he or she can get. Having a more accurate telescope to detect elusive objects always constitutes scientific progress.

The methodology we use in this paper — that is, harnessing a wide variety of cross-sectional anomalies to shed new light on an important problem in financial econometrics — is very much in tune with recent developments in other strands of the literature that are unrelated to covariance matrix estimation. For example, Hou et al. (2015) argue that the usefulness of a parsimonious model of expected stock returns should be judged against its ability to explain away a large number of cross-sectional anomalies. McLean and Pontiff (2016) measure the speed of convergence of financial markets towards informational efficiency by computing the decay rate of a large number of cross-sectional anomalies subsequent to academic publication.

Just as the merit for inventing DCC-NL does not belong to the present paper, the burden of proving that it is better than the multitude of covariance matrix estimators that have been proposed by countless authors does not fall on the present paper either. DCC-NL is the default choice at this juncture because it is the only one that addresses concomitantly the two major issues in the estimation of the covariance matrix of stock returns, namely conditional heteroskedasticy and the curse of dimensionality. Our point is only to establish that DCC-NL, as representative of best practices in covariance matrix estimation, has enough accuracy to reinstate the covariance matrix in its rightful place at the center of the Markowitz (1952) program and empirical asset pricing: The time has come to upgrade the practice of sorting into quantiles.

The paper is organized as follows. Section 2 gives a brief presentation of the DCC-NL covariance matrix estimator. Section 3 describes the empirical methodology for comparing test power with and without DCC-NL. Section 4 contains the empirical results. Section 6 concludes. Appendix A contains all tables and figures; Appendix B details the technique of

'Winsorization' that is applied to cross-sectional vectors of factors in our empirical work; and Appendix C details the set of factors we consider and how these factors are constructed in practice.

2 The DCC-NL Estimator of the Covariance Matrix

This brief recapitulation is only intended to make the present paper self-contained. The interested reader is referred to Engle et al. (2017) for the original exposition.

2.1 Time Variation in the Second Moments

The modelling and estimation of time-varying variances, covariances, and correlations requires aggregating the contributions from three different ideas.

2.1.1 Dynamic Conditional Correlation (DCC)

A key idea promoted by Engle (2002) is that modelling conditional heteroskedasticity is easy and successful in the univariate case, so we should take care of that *prior* to looking at the covariance matrix as a whole. Thus, for every asset $i = 1, \ldots, N$, we fit a GARCH(1,1) or similar model to the series i individually. Dividing the raw returns by the corresponding conditional standard deviations yields *devolatilized* returns that have unit variance. Call $\boldsymbol{s}_t$ the N-dimensional column vector of devolatilized residuals at time $t \in \{1, 2, \ldots, T\}$. Then the dynamics of the pseudo-correlation matrix Q_t can be specified as:

$$Q_t = \Theta + \alpha\, \boldsymbol{s}_{t-1}\boldsymbol{s}'_{t-1} + \beta\, Q_{t-1} \ , \tag{2.1}$$

where α and β are non-negative scalars satisfying $\alpha+\beta<1$ that govern the dynamics, and Θ is an N-dimensional symmetric positive definite matrix. Q_t is called a *pseudo*-correlation matrix because its diagonal terms are close, but not exactly equal, to one. Therefore the following adjustment is needed to recover the proper correlation matrix R_t:

$$R_t := \mathsf{Diag}(Q_t)^{-1/2}\, Q_t\, \mathsf{Diag}(Q_t)^{-1/2} \ , \tag{2.2}$$

where $\mathsf{Diag}(\cdot)$ denotes the function that sets to zero all the off-diagonal elements of a matrix.

2.1.2 Volatility Targeting

The second ingredient is the notion of "variance targeting" introduced by Engle and Mezrich (1996). Although originally invented in a univariate context, the extension to the multivariate case of interest here is straightfoward (Engle, 2002, Eq. (11)). The basic idea is that a suitable rescaling of the matrix Θ in equation (2.1) can be interpreted as the unconditional covariance matrix. Therefore, it can be estimated using standard techniques that ignore time series effects, separately from the other parameters. This approach yields the reparametrized model

$$Q_t = \Gamma(1-\alpha-\beta) + \alpha\, \boldsymbol{s}_{t-1}\boldsymbol{s}'_{t-1} + \beta\, Q_{t-1} \ , \tag{2.3}$$

where Γ is the long-run covariance matrix of the devolatilized returns $\boldsymbol{s}_t$ for $t = 1, \ldots, T$.[1]

2.1.3 Composite Likelihood

After having dealt with the conditional variances and partialled out the problem of estimating the unconditional covariance matrix, the only remaining task is to estimate the dynamic correlation parameters α and β. These two scalars play the same role as their counterparts in the more familiar ARMA(1,1) and GARCH(1,1) models, but for conditional correlation matrices.

When the matrix dimension is large, say $N = 1000$, the standard likelihood maximization technique would require inverting T matrices of dimension 1000×1000 at every iteration, which is numerically challenging. Pakel et al. (2017) found a more efficient solution called the 2MSCLE method: combine the individual likelihoods generated by 2×2 blocks of contiguous variables. Maximizing this *composite likelihood* yields asymptotically consistent estimators for α and β, as long as the DCC model is well-specified. The intuition is that every individual correlation coefficient shows traces of the dynamic parameters α and β in its own time series evolution, so a sufficiently large subset of individual correlations will reveal (a consistent approximation of) the true parameters. The advantage of this procedure is that it is numerically stable and fast in high dimensions; for example, Engle et al. (2017) manage to take it to a large universe of $N = 1000$ stocks.

2.1.4 DCC Estimation Procedure

To summarize, the estimation of the DCC model unfolds in three steps:

1. Fit a univariate GARCH(1,1) model to every stock return series individually, and divide the raw returns by their conditional standard deviations to devolatilize them.
2. Estimate the unconditional covariance matrix of devolatilized returns somehow.
3. Maximize the 2MSCLE composite likelihood to obtain consistent estimators of the two parameters of correlation dynamics in a numerically stable and efficient way.

At this juncture, it becomes apparent from step 2 that we need an estimator of the unconditional covariance matrix of devolatilized returns that performs well when the dimension is large.[2]

2.2 Estimation of Large-Dimensional Unconditional Covariance Matrices

The reader is invited to peruse Ledoit and Wolf (2012, 2015) for a more detailed treatment.

[1]Since the devolatilized returns all have unit variance, Γ is actually a proper correlation matrix, that is, its diagonal elements are all equal to one.

[2]Note that in practice the devolatilized returns have to be based on estimated univariate GARCH models rather than the 'true', unobservable univariate GARCH models.

2.2.1 Spectral Decomposition

The textbook estimator of $\mathit{\Gamma}$ is the sample covariance matrix $C := \sum_{t=1}^{T} \boldsymbol{s}_t \boldsymbol{s}_t'/T$. Both matrices admit spectral decompositions:

$$C = \sum_{i=1}^{N} \lambda_i \cdot \boldsymbol{u}_i \boldsymbol{u}_i' \qquad \text{and} \qquad \mathit{\Gamma} = \sum_{i=1}^{N} \tau_i \cdot \boldsymbol{v}_i \boldsymbol{v}_i' \ , \tag{2.4}$$

where $(\lambda_1, \ldots, \lambda_N; \boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$ denotes a system of eigenvalues and eigenvectors of the sample covariance matrix C, and $(\tau_1, \ldots, \tau_N; \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N)$ denotes a system of eigenvalues and eigenvectors of the population covariance matrix $\mathit{\Gamma}$. Eigenvalues are indexed in ascending order without loss of generality.

In the traditional asymptotic framework, where the sample size T goes to infinity, while the number of assets N remains constant, the sample eigenvalue λ_i is a consistent estimator of its population counterpart τ_i, and the sample eigenvector $\boldsymbol{u}_i$ is a consistent estimator of its population counterpart $\boldsymbol{v}_i$, for $i = 1, \ldots, N$. However, this asymptotic framework is not robust against the curse of dimensionality. When N is no longer negligible with respect to T, the sample spectrum is far from its population counterpart.

This is why it is necessary to turn to another asymptotic framework that offers a different family of analytical solutions. Unlike the formulas from traditional asymptotics, they work also if N is not negligible with respect to T, and even if N is greater than T. The key assumption is that the ratio N/T converges to some limit $c \in [0, +\infty)$ called the *concentration (ratio)*. This framework is called *large-dimensional* asymptotics, and it includes traditional (fixed-dimensional) asymptotics as a special case when the concentration c is equal to zero. Thus, it is a generalization of traditional asymptotics that is able to cope with the curse of dimensionality by making necessary corrections (whose intensity increases in c) to the standard formulas.

2.2.2 Portfolio Selection

Stein (1986) argued that, in the absence of *a priori* knowledge about the structure of the eigenvectors of the (unobservable) population covariance matrix $\mathit{\Gamma}$, estimators should preserve the sample covariance matrix eigenvectors $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$, and correct the sample eigenvalues only. This framework is called *rotation-equivariant* because the economic outcome is immune to repackaging the N original stocks into a collection of N funds investing in these stocks, as long as the funds span the same investment universe as the stocks.

It is easy to show that, among rotation-equivariant estimators of the covariance matrix, the one that performs the best across all possible linear constraints for the purpose of portfolio selection in terms of minimizing out-of-sample variance is:

$$\widetilde{C} := \sum_{i=1}^{N} \underbrace{\left(\boldsymbol{u}_i' \mathit{\Gamma} \boldsymbol{u}_i\right)}_{\phi_i} \cdot \boldsymbol{u}_i \boldsymbol{u}_i' \ . \tag{2.5}$$

This makes economic sense because $\boldsymbol{u}_i' \Gamma \boldsymbol{u}_i$ is the out-of-sample variance of the portfolio whose weights are given by the ith sample eigenvector $\boldsymbol{u}_i$. Thus we notice the emergence of a third quantity, after the sample eigenvalue $\lambda_i = \boldsymbol{u}_i' C \boldsymbol{u}_i$, and the population eigenvalue $\tau_i = \boldsymbol{v}_i' \Gamma \boldsymbol{v}_i$: the hybrid $\phi_i := \boldsymbol{u}_i' \Gamma \boldsymbol{u}_i$, which represents the best we can do with the sample eigenvectors.

The key is that, under large-dimensional asymptotics, the vectors $\boldsymbol{\lambda} := (\lambda_i)_{i=1,\dots,N}$, $\boldsymbol{\tau} := (\tau_i)_{i=1,\dots,N}$, and $\boldsymbol{\phi} := (\phi_i)_{i=1,\dots,N}$ are all far apart from one another. It is only as the concentration c goes to zero, that is, as we approach standard (fixed-dimension) asymptotics, that their mutual differences vanish. When $c > 0$, which is the case when the investment universe is large, appropriate corrections must be applied to go from $\boldsymbol{\lambda}$ to $\boldsymbol{\tau}$ to $\boldsymbol{\phi}$.[3] Qualitatively, $\boldsymbol{\lambda}$, $\boldsymbol{\tau}$, and $\boldsymbol{\phi}$ have the same cross-sectional average, but $\boldsymbol{\lambda}$ is more dispersed than $\boldsymbol{\tau}$, which in turn is more dispersed than $\boldsymbol{\phi}$.

2.2.3 NonLinear (NL) Shrinkage Estimator of the Covariance Matrix

The ideal would be to have two deterministic functions $\boldsymbol{\Lambda}^{N,T}$ and $\boldsymbol{\Phi}^{N,T}$ from $[0,+\infty)^N$ to $[0,+\infty)^N$ mapping out the two important expectations:

$$\boldsymbol{\tau} \longmapsto \boldsymbol{\Lambda}^{N,T}(\boldsymbol{\tau}) := \big(\Lambda_1^{N,T}(\boldsymbol{\tau}),\dots,\Lambda_N^{N,T}(\boldsymbol{\tau})\big) = \big(\mathbb{E}[\lambda_1],\dots,\mathbb{E}[\lambda_N]\big) = \big(\mathbb{E}[\boldsymbol{u}_1' C \boldsymbol{u}_1],\dots,\mathbb{E}[\boldsymbol{u}_N' C \boldsymbol{u}_N]\big)$$
$$\boldsymbol{\tau} \longmapsto \boldsymbol{\Phi}^{N,T}(\boldsymbol{\tau}) := \big(\Phi_1^{N,T}(\boldsymbol{\tau}),\dots,\Phi_N^{N,T}(\boldsymbol{\tau})\big) = \big(\mathbb{E}[\phi_1],\dots,\mathbb{E}[\phi_N]\big) = \big(\mathbb{E}[\boldsymbol{u}_1' \Gamma \boldsymbol{u}_1],\dots,\mathbb{E}[\boldsymbol{u}_N' \Gamma \boldsymbol{u}_N]\big)$$

Then we would use the observed eigenvalues of the sample covariance matrix, $\boldsymbol{\lambda}$, to reverse-engineer an estimator of the population eigenvalues by solving the optimization problem

$$\widehat{\boldsymbol{\tau}} := \underset{\mathbf{t}\in[0,+\infty)^N}{\operatorname{argmin}} \frac{1}{N}\sum_{i=1}^{N}\Big(\Lambda_i^{N,T}(\mathbf{t}) - \lambda_i\Big)^2 , \tag{2.6}$$

and the nonlinear shrinkage estimator of the covariance matrix would follow as

$$\widehat{C} := \sum_{i-1}^{N} \Phi_i^{N,T}(\widehat{\boldsymbol{\tau}}) \cdot \boldsymbol{u}_i \boldsymbol{u}_i' . \tag{2.7}$$

Due to tractability issues, however, we only know approximations to the functions $\boldsymbol{\Lambda}^{N,T}$ and $\boldsymbol{\Phi}^{N,T}$ that are valid asymptotically as the universe dimension N goes to infinity along with the sample size T, with their ratio N/T converging to the concentration c. Ledoit and Wolf (2012, 2015) show that replacing the true expectation functions with their approximations can be done at no loss asymptotically. Therefore, this procedure yields a nonlinear shrinkage estimator of the covariance matrix that is optimal in the large-dimensional asymptotic limit.

Qualitatively speaking, the effect of composing $\boldsymbol{\Phi}^{N,T}$ with the inverse of $\boldsymbol{\Lambda}^{N,T}$ (or approximations thereof) moves the sample eigenvalues closer to one another, while preserving their cross-sectional average. The effect is increasing in N/T and highly nonlinear; for example, isolated eigenvalues that lie near the bulk of the other eigenvalues move in the direction of the bulk more than those distant from the bulk.

[3] Correcting these relationships when the ratio of variables to observations is significant is analogous to correcting Newtonian relationships when the ratio of velocity to speed of light is significant (Einstein, 1905).

2.3 DCC-NL Model

In summary, the estimation of the DCC-NL model of Engle et al. (2017) proceeds as follows:

1. Fit univariate GARCH models to devolatilize returns.
2. Compute the sample covariance matrix of devolatilized returns.
3. Decompose it into eigenvalues and eigenvectors.
4. Invert an approximation of the function $\boldsymbol{\Lambda}^{N,T}$ to estimate population eigenvalues.
5. Apply an approximation of the function $\boldsymbol{\Phi}^{N,T}$ to shrink eigenvalues nonlinearly.
6. Recompose with the sample eigenvectors to estimate the unconditional covariance matrix Γ in (2.3).
7. Transform the resulting estimator of Γ from a covariance matrix to a proper correlation matrix.[4]
8. Maximize the 2MSCLE composite likelihood to estimate the correlation dynamics.
9. Recombine the estimated conditional correlation matrix with the estimated univariate GARCH processes to obtain an estimated conditional covariance matrix.

The outside steps (1–2 and 7–9) compose the DCC part, while the inside steps (3–6) compose the NL part of the DCC-NL estimation procedure. The final product is a time series of N-dimensional conditional covariance matrix estimates, which we call $\{H_t\}_{t=1}^{T}$. More explicit formulas are provided in Engle et al. (2017).

3 Empirical Methodology

The goal is to construct long-short portfolios exposed to a given factor. The size of the investment universe is denoted by N, and stocks in this universe are indexed by i. Days on which trading takes place are indexed by t. The cross-sectional vector of factor scores observable at the beginning of day t is denoted by $\boldsymbol{m}_t := (m_{t,1}, \ldots, m_{t,N})'$. A portfolio of the type that we consider is defined by a weight vector $\boldsymbol{w}_t := (w_{t,1}, \ldots, w_{t,N})'$ that satisfies

$$\sum_{w_{t,i}<0} |w_{t,i}| = \sum_{w_{t,i}>0} |w_{t,i}| = 1 \ .$$

Note that the weights of such a long-short portfolio necessarily sum to zero, that is, the portfolio is dollar-neutral (on the day of portfolio construction). Furthermore, the gross exposure of the portfolio is two dollars (on the day of portfolio construction).

3.1 Portfolios: Quantile-Based Weighting

Let B be the number of quantiles considered; for example, $B = 3$ for terciles, $B = 5$ for quintiles, and $B = 10$ for deciles. Let d be the largest integer that is smaller than or equal

[4]Doing so is motivated by the fact that Γ itself is a proper correlation matrix, as pointed out previously.

to N/B. Finally, let $\{(1), (2), \ldots, (N)\}$ be permutation of $\{1, 2, \ldots, N\}$ that results in ordered factor scores (from smallest to largest):

$$m_{t,(1)} \leq m_{t,(2)} \leq \ldots \leq m_{t,(N)} .$$

Then the quantile-based portfolio is given by the weight vector $\boldsymbol{w}_t^{\text{Qu}}$ with

$$\begin{aligned} w_{t,(1)}^{\text{Qu}} = \ldots = w_{t,(d)}^{\text{Qu}} &:= -1/d , \\ w_{t,(d+1)}^{\text{Qu}} = \ldots = w_{t,(N-d)}^{\text{Qu}} &:= 0 , \text{ and} \\ w_{t,(N-d+1)}^{\text{Qu}} = \ldots = w_{t,(N)}^{\text{Qu}} &:= 1/d . \end{aligned}$$

The resulting portfolio return on such a portfolio-construction day is denoted by $r_t^{\text{Qu}} := \boldsymbol{x}_t' \boldsymbol{w}_t^{\text{Qu}}$, where $\boldsymbol{x}_t$ is the $N \times 1$ vector of returns at date t. In case the portfolio is not updated every day, it is customary to hold number of shares rather than portfolio weights constant until the next portfolio construction. (Holding portfolio weights constant would require daily rebalancing, which would incur additional trading costs.) In this case, the portfolio return r_t^{Qu} on a given day depends on how the vector of portfolio weights has 'evolved' over time due to the price movements of the various stocks in the portfolio.[5]

3.2 Portfolios: Efficient Weighting

The alternative investment problem we propose is formulated as

$$\min_{\boldsymbol{w}} \boldsymbol{w}' H_t \boldsymbol{w} \tag{3.1}$$

$$\text{subject to} \quad \boldsymbol{m}_t' \boldsymbol{w} = \boldsymbol{m}_t' \boldsymbol{w}_t^{\text{Qu}} \quad \text{and} \tag{3.2}$$

$$\sum_{w_i < 0} |w_i| = \sum_{w_i > 0} |w_i| = 1 , \tag{3.3}$$

where H_t is the DCC-NL estimator of the covariance matrix of $\boldsymbol{x}_t$. Denote a solution of this investment problem by $\boldsymbol{w}_t^{\text{Ef}}$. The resulting portfolio return on such a portfolio-construction day is denoted by $r_t^{\text{Ef}} := \boldsymbol{x}_t' \boldsymbol{w}_t^{\text{Ef}}$. In case the portfolio is not updated every day, it is customary to hold number of shares rather than portfolio weights constant until the next portfolio construction; in this case, the portfolio return r_t^{Ef} on a given day depends on how the vector of portfolio weights has 'evolved' over time due to the price movements of the various stocks in the portfolio.

The motivation here is that we want to construct a portfolio that (i) has the same exposure to the vector of factors $\boldsymbol{m}_t$ as the quantile-based sorting portfolio because of (3.2), but (ii) has a smaller variance because of (3.1). If this goal is accomplished, then the resulting portfolio returns will generally result in a larger (in magnitude) 'Student' t-statistic (cf. (3.10) below), since the smaller variance of the returns will result in a smaller standard error in the denominator of the t-statistic whereas the sample average in numerator will be roughly the same. This means the power of the test should increase. It is key to have an accurate estimate

[5] In particular, the portfolio will no longer be necessarily dollar-neutral until the next portfolio construction.

of the covariance matrix of $\boldsymbol{x}_t$ in order to achieve this goal: this where the DCC-NL model comes in.

Remark 3.1 (Expected Portfolio Returns)**.** Since the factor $\boldsymbol{m}_t$ is only a proxy for the first moment of returns, it cannot be guaranteed via (3.2) that the two portfolios have the same expected return. But the two portfolios certainly have the same factor exposure and their expected returns should be close to each other, with neither being systematically higher or lower than the other. Consequently, any systematic improvement in the efficient-weighting portfolio compared to the quantile-based portfolio can be attributed to the reduction of the portfolio variance. It has been argued that Markowitz-type portfolios implicitly pick up risk-based pricing anomalies, leading to higher Sharpe ratios; for example, see Haugen and Baker (1991) and Scherer (2011). But such arguments only apply to portfolios that are fully invested and not to dollar-neutral long-short portfolios, as we consider in this paper: Any 'systematic' effect based on picking up risk-based pricing anomalies would apply both to the long portfolio and to the short portfolio and thus would cancel out in the final long-short portfolio. ■

Remark 3.2 (Estimator of the Covariance Matrix)**.** The efficient-weighting portfolio is not necessarily tied to the DCC-NL model: One could use other estimators of the covariance matrix of $\boldsymbol{x}_t$ in (3.1) instead, such as the linear shrinkage estimators of Ledoit and Wolf (2003, 2004a,b) or the nonlinear shrinkage estimator of Ledoit and Wolf (2017). Note that all these estimators are static estimators, that is, unlike the dynamic DCC-NL estimator, they do not incorporate the time-varying nature of the covariance matrix of $\boldsymbol{x}_t$.

In the empirical analysis of Section 4, we update the portfolios only once per month, which is standard in the finance literature. In such a scheme, static estimators of the covariance matrix might work similarly well as a (sophisticated) dynamic estimator. But our methodology should also be of interest to real-life portfolio managers. Such managers generally update their portfolios on a daily basis, which certainly motivates the use of a (sophisticated) dynamic estimator of the covariance matrix.

Theoretical reasons for using the DCC-NL method include: (i) it is state-of-the-art; (ii) ARCH/GARCH effects are real and have been known for decades, so ignoring them is hard to justify; (iii) any estimation method overweights the recent past and underweights (or zero-weights) the distant past, and the exponential-type decay implicit in a GARCH model is a more meaningful way to do that smoothly compared to an 'awkward' choice of the (relatively short) window length for a static model, that is, the length of the period where all the observations receive equal weight before the weights drop to zero in a discontinuous fashion.

For all these reasons, we promote the use of the DCC-NL estimator.[6] ■

[6] An extensive comparison to other estimators of the covariance matrix in the empirical analysis of Section 4 would have been computationally prohibitive.

3.3 Connection with Markowitz Portfolio Selection

There is a direct connection with the well-known mean-variance portfolio optimization problem for dollar-neutral long-short strategies:

$$\min_{\boldsymbol{w}} \boldsymbol{w}' H_t \boldsymbol{w} \tag{3.4}$$

$$\text{subject to} \quad \boldsymbol{m}_t' \boldsymbol{w} = \gamma_t \quad \text{and} \tag{3.5}$$

$$\sum_{w_i<0} |w_i| = \sum_{w_i>0} |w_i| \ , \tag{3.6}$$

where γ_t denotes some target exposure to the factor $\boldsymbol{m}_t$. The formulation in (3.6) is traditionally expressed as $\sum_i w_i = 0$, which is mathematically equivalent. We only state it in the shape of (3.6) to emphasize the analogy with (3.3). The difference is that (3.3) constrains the gross portfolio size to be equal to that of the long-short portfolio based on sorting, which is two, in order to make the two strategies directly comparable. The general solution to (3.4)–(3.6) is of the form

$$\boldsymbol{w}_t = \kappa_t \cdot \boldsymbol{w}_t^{\mathrm{MV}} \ , \quad \text{where} \quad \boldsymbol{w}_t^{\mathrm{MV}} := \underbrace{\frac{H_t \boldsymbol{m}_t}{\mathbb{1}' H_t \boldsymbol{m}_t}}_{\text{tangency portfolio}} - \underbrace{\frac{H_t \mathbb{1}}{\mathbb{1}' H_t \mathbb{1}}}_{\text{global minimum variance portfolio}} \ . \tag{3.7}$$

In (3.7), $\mathbb{1}$ denotes the unit vector of dimension N, κ_t denotes some scaling parameter proportional to the target exposure γ_t, and the superscript MV stands for mean-variance optimization. Mathematically speaking, the scaling factor κ_t is strictly positive if and only if the target exposure γ_t is itself strictly positive (except in the degenerate case where the factor $\boldsymbol{m}_t$ is proportional to $\mathbb{1}$); this is the only economically relevant case, as we seek positive exposure to any candidate factor.

In plain terms, the minimum-variance weights require going long one dollar in the tangency portfolio and short one dollar in the global minimum variance portfolio, up to some multiplier κ_t. This is not exactly what we do in $\boldsymbol{w}_t^{\mathrm{Ef}}$ because our purpose is to make an apples-to-apples comparison with sorting, but on a stand-alone basis (3.7) may make more sense. As for the scaling parameter κ_t, there are two obvious proposals:

- $\boldsymbol{w}_t^{\mathrm{MV}}$: Take $\kappa_t = 1$ so that there is ± 1 dollar in the two 'basis' portfolios: the tangency portfolio and the global minimum variance portfolio.
- $\boldsymbol{w}_t^{|\mathrm{MV}|}$: Choose the unique $\kappa_t > 0$ that satisfies (3.3), so that there is ± 1 dollar in the combined portfolio. Thus,

  $$\boldsymbol{w}_t^{|\mathrm{MV}|} = 2 \, \frac{\boldsymbol{w}_t^{\mathrm{MV}}}{\left\| \boldsymbol{w}_t^{\mathrm{MV}} \right\|_1} \ , \tag{3.8}$$

  where $\| \cdot \|_1$ denotes the L_1-norm of a vector.

The gross exposure of $\boldsymbol{w}_t^{|\mathrm{MV}|}$ is two dollars by construction, whereas the gross exposure of $\boldsymbol{w}_t^{\mathrm{MV}}$ can be anything in principle. Due to the fact that the offset between the long and short positions in the tangency portfolio and the global minimum variance portfolio varies through time with probability one, there will be a difference in the resulting t-statistics, but it is hard to ascertain *ex ante* which one is to be favored. If anything, the first one, being based on the matrix algebra of mean-variance optimization, has more appeal to financial econometricians; whereas the second one makes more sense for long-short market-neutral hedge fund managers whose prime brokers limit gross leverage. What $\boldsymbol{w}_t^{\mathrm{MV}}$ and $\boldsymbol{w}_t^{|\mathrm{MV}|}$ have in common is that they are both less constrained than the estimator we focus on for the purpose of an apples-to-apples comparison with quantile-based weighting, namely, $\boldsymbol{w}_t^{\mathrm{Ef}}$.

Other choices for the scaling parameter κ_t are possible, such as the inverse of the L_2-norm of the original weight vector $\boldsymbol{w}_t^{\mathrm{MV}}$, a volatility budget, etc. This opens up the door to a whole new family of variance-minimizing factor portfolio weighting schemes. A full investigation of this line of research, however, lies well outside the scope of the present paper because we only aim to establish that a sophisticated estimator of the (inverse) covariance matrix boosts the power of predictive tests for cross-sectional anomalies relative to common, quantile-based portfolios.

3.4 Tests for Predictive Ability

The ability of a factor to forecast the cross-section of stock returns is judged by whether a long-short portfolio exploiting the factor can deliver returns with a positive expected value. In particular, we consider the hypothesis testing problem

$$H_0 : \mathbb{E}(r_t^{\mathrm{St}}) \leq 0 \quad \text{vs.} \quad H_1 : \mathbb{E}(r_t^{\mathrm{St}}) > 0 \ , \tag{3.9}$$

where St $\in$ {Qu,Ef} stands for one of the two strategies, quantile-based weighting or efficient weighting

The test is based on observed strategy returns $r_t^{\mathrm{St}}, t = 1, \ldots, T$. The 'Student' t-statistic of the test is given by

$$t^{\mathrm{St}} := \frac{\bar{r}^{\mathrm{St}}}{\mathrm{SE}(\bar{r}^{\mathrm{St}})} \quad \text{with} \quad \bar{r}^{\mathrm{St}} := \frac{1}{T}\sum_{t=1}^{T} r_t^{\mathrm{St}} \ , \tag{3.10}$$

where $\mathrm{SE}(\bar{r}^{\mathrm{St}})$ denotes a standard error of $\bar{r}^{\mathrm{St}}$. The common choice in the literature for such a standard error is the 'naïve' standard error based on an assumption of independent and identically distributed (i.i.d.) returns. Specifically, it is given by $s^{\mathrm{St}}/\sqrt{T}$, where s^{St} denotes the sample standard deviation of the observed returns $r_t^{\mathrm{St}}, t = 1, \ldots, T$.

Instead, we consider it important to use a HAC standard error that is robust against heteroskedasticity and serial correlation in the returns. In particular, we use the standard error based on the quadratic spectral (QS) kernel with automatic choice of bandwidth as detailed in Andrews (1991).

The common critical value in the literature is two: If the t-statistic is larger than two, the factor is deemed successful. On the other hand, Harvey et al. (2016) call for a more demanding critical value of three due to multiple-testing issues.

4 Empirical Analysis

4.1 Data and General Portfolio-Construction Rules

We download daily stock return data from the Center for Research in Security Prices (CRSP) starting in 01/01/1980 and ending in 12/31/2015. We restrict attention to stocks from the NYSE, AMEX, and NASDAQ stock exchanges.

For simplicity, we adopt the common convention that 21 consecutive trading days constitute one 'month'. The out-of-sample period ranges from 01/08/1986 through 12/31/2015, resulting in a total of 360 'months' (or 7560 days). All portfolios are updated monthly.[7] We denote the investment dates by $h = 1, \ldots, 360$. At any investment date h, the efficient-weighting portfolio (3.1)–(3.3) uses the DCC-NL estimate H_t of the covariance matrix based on the most recent 1250 daily returns, which roughly corresponds to using five years of past data. The portfolio based on sorting uses quintiles, which seems to be the most common choice in the literature.

We consider the following portfolio sizes: $N \in \{100, 500, 1000\}$. For a given combination (h, N), the investment universe is obtained as follows. We find the set of stocks that have a complete return history over the most recent $T = 1250$ days as well as a complete return 'future' over the next 21 days.[8] We then look for possible pairs of highly correlated stocks, that is, pairs of stocks that have returns with a sample correlation exceeding 0.95 over the past 1250 days. With such pairs, if they should exist, we remove the stock with the lower volume of the two on investment date h.[9] Of the remaining set of stocks, we then pick the largest N stocks (as measured by their market capitalization on investment date h) as our investment universe. In this way, the investment universe changes slowly from one investment date to the next.

We consider a total of 62 factors taken from Green et al. (2013) and Hou et al. (2015); the corresponding data are downloaded from the merged CRSP/Compustat database. Table 1 lists the factors and Appendix C contains a detailed description of how the factor scores are computed. Note that for $N = 1000$, there are not sufficient data available for factors 23, 32, 37, and 52–57. We apply 'Winsorizaton' to any cross-sectional vector of factor scores $\boldsymbol{m}_t$ in order to migitate potential problems with 'outlying' scores that are unusually large in magnitude; see Appendix B for the corresponding details.

4.2 Main Results

The individual t-statistics are detailed in Table 2. Not surprisingly, in some cases the t-statistic for quantile-based weighting (though generally not significantly so). It can be

[7]monthly updating is common practice to avoid an unreasonable amount of turnover and thus transaction costs. During a month, from one day to the next, we hold number of shares fixed rather than portfolio weights; in this way, there are no transactions at all during a month.

[8]The latter, 'forward-looking' restriction is not a feasible one in real life but is commonly applied in the related finance literature on the out-of-sample evaluation of portfolios.

[9]The reason is that we do not want to include highly similar stocks; in the early years, there are no such pairs; in the most recent years, there are never more than three such pairs.

assumed that the corresponding factors will be discarded immediately by a researcher, since they can never be established as successful based on a negative t-statistic. For each universe size $N \in \{100, 500, 1000\}$, we therefore restrict attention to factors for which quantile-based weighting yield a positive t-statistic. For such factors, we also present the value of the ratio of the two t-statistics: the one for efficient weighting divided by the one for quantile-based weighting.

Table 3 presents the average ratio for each universe size $N \in \{100, 500, 1000\}$. The average ratio is always larger than two, meaning that, on average, the t-statistic more than doubles when a researcher upgrades from quantile-based weighting to efficient weighting.

It is natural to ask whether these averages might be influenced by a few 'outlying' ratios which can occur when the t-statistic for quantile-based weighting (which appears in denominator) is close to zero. For example, take the case of factor 33 with a universe size $N = 100$. In this case, the t-statistic for quantile-based weighting equals 0.020 whereas the t-statistic for efficient weighting equals 1.048, resulting in a ratio of 52.4. Consequently, we also compute averages only for cases where the t-statistic for quantile-based weighting is bounded away from zero. First, we only consider cases where the t-statistic is is larger than 0.5; second, we only consider cases where the t-statistic is larger than 1.0. The corresponding averages are also found in Table 3. It can be seen that the averages decrease as the lower bound increases (from 0 to 0.5 to 1.0), especially for $N = 100$. But when the lower bound is 0.5, the averages for $N = 500, 1000$ still exceed two; and when the lower bound is 1.0, the averages for $N = 500, 1000$ are still close to two (if less than two now). Therefore, the impressive power gains of efficient weighting over quantile-based weighting are not solely driven by a few t-statistics for quantile-based weighting that are close to zero.

Arguably, it is of main interest how much the number (and proportion) of significant factors increase when a researcher upgrades from quantile-based weighting to efficient weighting. The common critical value in the literature for the value of a t-statistic is two. On the other hand, Harvey et al. (2016) argue that a critical value of three should be used instead because of multiple-testing issues. We consider both critical values, two and three, in Table 4. One can see that for both strategies, Qu(antile-based weighting) and Ef(ficient weighting), the proportion of significant factors increases in N;[10] therefore, it is in the best interest of researchers to use as large an investment universe as possible. One can further see that the numbers/proportions of significant factors are always much larger for efficient weighting compared to quantile-based weighting. In particular, when a critical value of three is used, the number/proportion of significant factors more than doubles for all universe sizes when a researcher upgrades to efficient weighting.

[10]This is not always true for the *number* of significant factors, which is due to the fact that we only have 53 factors available for $N = 1000$ compared to 62 factors for $N \in \{100, 500\}$, as detailed in Section C.

5 Further Results and Related Literature

5.1 Robustness Check: Filtering through Fama-French Factors

Instead of focusing on the expected value of the portfolio returns, we now shift focus to the intercept (alpha) of a regression of the portfolio returns on the five factors of Fama and French (2015). The numerator of the t-statistic now is the ordinary least squares estimator of the intercept of this regression and the denominator of the t-statistics now is the corresponding HAC standard error.[11]

The results are presented in Tables 5 and 6.[12] It can be seen that the relative power gains of efficient weighting over quantile-based weighting are somewhat reduced. (In particular, the proportion of significant factors compared to using the raw portfolio returns generally increases for quantile-based weighting whereas it generally decreases for efficient weighting.) Nevertheless, efficient weighting is still generally more powerful than quantile-based weighting.[13]

5.2 Spanning Test

It would be interesting to find out where the empirical advantage of the efficient-weighting procedure comes from: Is it because the procedure selects stocks that are more informative or because it minimizes diversifiable — and hence unpriced — risks? One way to distinguish between the two possible explanations is to check whether the efficient factors span conventional portfolios, such as the 25 Fama-French size and value portfolios, in the spirit of Welch (2008, Section IV).

Table 7 shows that the conventional, quantile-based factors do not span the 25 Fama-French size and value portfolios, which extends the results of Welch (2008) from 6 size and value portfolios to 25. The same is true of the efficient-weighting factors. Therefore, the empirical advantage of our new proposed procedure does not come from its ability to select more informative stocks; it must come instead from its ability to minimize exposure to unpriced risk.

5.3 Summary Statistics of Portfolio Weights

We compute some summary statistics on the vectors of portfolio weights $\boldsymbol{w}_t^{\text{St}}$ over time, for $\text{St} \in \{\text{Qu,Ef}\}$. In each month, we compute the following four characteristics:

- **Min:** Minimum weight
- **Max:** Maximum weight
- **SD**: Standard deviation of weights

[11]We again use the QS kernel with automatic choice of bandwidth as detailed in Andrews (1991).

[12]We also carried out the analysis by using the traditional three factors of Fama and French (1992) for filtering and the results are qualitatively similar.

[13]The only exception is the case $N = 100$ with a critical value of three.

- **MAD:** Mean absolute deviation of Ef portfolio from Qu portfolio computed as

$$\frac{1}{N}\sum_{i=1}^{N}\left|w_{t,i}^{\mathrm{Ef}} - w_{t,i}^{\mathrm{Qu}}\right| .$$

For each characteristic, we then compute the average outcome over the 360 portfolio formations (that is, over the 360 'months') and summarize the resulting numbers — 62 numbers for $N \in \{100, 500\}$ and 53 numbers for $N = 1000$ — by means of a boxplot. (Note that for the first three characteristics, the summary statistic for the Qu portfolio is constant over all the factors under consideration and so the boxplot really involves the summary statistics for the Ef portfolio only, apart from a dashed line which indicates the corresponding, constant summary statistic for the Qu portfolio.) These boxplots are displayed in Figures 1–3.

Given that the Qf portfolios have uniform weights, clearly the mininum and maximum weights of the Ef portfolios will extend further out; the question is by how much? In the case $N = 100$, they are typically three times larger in absolute value, that is, a $\pm 5\%$ minimum/maximum weight becomes of the order of $\pm 15\%$. A multiplicative factor of three is relatively acceptable, given that we make no effort to control extreme weights, and that quantile-based weights are the tamest by construction. In the case $N = 1000$, the extreme weights of the Ef portfolios are even smaller in magnitude, typically on the order of $\pm 4\%$ only, which is reasonable for a well-diversified strategy. The case $N = 500$ is essentially an interpolation between the other two dimensions.

In addition, we provide boxplots for average turnover[14] for any given combination (St, N). These boxplots are displayed in Figure 4; note that for both Qu and Ef a number of 1 corresponds to a turnover of 100% per month. This scale may appear high on paper, but in practice it is fairly representative of the volume executed by Statistical Arbitrage hedge funds that specialize in systematically exploiting large ensembles of cross-sectional anomalies: Their automated order-placing system usually brings the transaction cost down to a range as low as 3–5 basis points.

Similarly to extreme positions, it is to be expected *a priori* that the Ef portfolios will turn over more than the Qu portfolios; the question, again, is by how much? Median turnover typically goes up by a multiplicative factor of less than 1.5, which is in the reasonable zone, given that we make no effort to control turnover. This statistic implies that most of the turnover is due to changes in the factor scores themselves, rather to the covariance matrix. Interestingly, turnover decreases as the number of stocks in the investment universe, N, increases, for both the Qu and Ef portfolios.

In summary, not only does power increase in N, as described in Section 4.2, but also maximum absolute weights and turnover decreases in N: It is thus in everyone's interest to use a large investment universe, indeed.

[14] The average is taken over the 359 turnovers from the end of a given month to the start of the next month.

5.4 Related Literature

A recent working paper by Cattaneo et al. (2018) tries to improve upon the common choice of sorting into five or ten portfolios, that is, using quintiles or deciles, by developing a data-driven procedure to select the optimal number of portfolios. They merely modulate the sorting mechanism, whereas we shift completely away of this paradigm by playing the covariance-matrix card.

The general thrust of their paper is to slice up the universe using more than ten quantiles, and pitch a very narrow selection of the top-scoring stocks against a very narrow selection of the bottom-scoring stocks. This approach may work well statistically. But from the point of view of investability, it focuses a lot of the trading volume onto such a small number of stocks that it may not be implementable in practice without dire price impact for institutional asset sizes. As Korajczyk and Sadka (2004) argue, exploitability is a key feature of an anomaly.

Cattaneo et al. (2018) also estimate nonlinearities in the mapping from factor score to expected returns, whereas we just work within the much simpler linear framework. They find the size effect is convex and momentum concave. However, in order to identify the shape, they must first data-snoop (cf. Lo and MacKinlay (1990)) the whole sample, which would make any test utilizing such knowledge non-predictive.

Another recent related working paper is by Daniel et al. (2017). The major commonality is their equation (18), which is quite close to the way we construct the efficient portfolio weights $\boldsymbol{w}_t^{\mathrm{Ef}}$, up to suitable rescaling. However, this is not the approach that they use in practice because, as they themselves acknowledge, "there are well-known issues associated with estimating [the covariance matrix of a large universe of stocks] and using it to do portfolio formation." As a result, they develop an alternative approach that only uses covariances between a small set of factor returns (up to five), which is completely different from our own approach. Indeed, one of our major contributions is to solve the well-known issues associated with estimating large-dimensional covariance matrices by employing the DCC-NL model.

6 Conclusion

This paper demonstrates that, in accordance with the theory of mean-variance optimization, portfolio construction in predictive tests of cross-sectional anomalies should incorporate a suitable estimator of the covariance matrix of stock returns. When a researcher upgrades from quantile-based weighting to efficient weighting based on the DCC-NL covariance matrix estimator of Engle et al. (2017), 'Student' t-statistics, on average, more than *double* — across a large panel of return-predictive signals (or "factors") — when the investment universe is large. This power boost is especially needed because multiple-testing issues may justify raising the t-statistic significance threshold from its usual level of two to a more demanding level of three, as proposed by Harvey et al. (2016). The power boost also cures the inherent handicap of short-history datasets by multiplying the effective number of years by approximately four in large dimensions. Cross-sectional testing methodologies that do not use a suitable estimator of the covariance matrix, such as DCC-NL, are underpowered and their use should be discouraged.

Directions for further research include (i) exploring the performance of the alternative, Markowitz portfolios described in Section 3.3; (ii) using more accurate univariate models than the straightforward GARCH(1,1) to devolatilize individual return series in the first step of the procedure (such as models that incorporate asymmetric responses and/or intraday prices); (iii) pre-conditioning the cross-section of stock returns by a low-dimensional model with exogenous risk factors; and (iv) using the inverse of the DCC-NL covariance matrix to construct portfolios that would yield a more efficient test of an asset-pricing model, in the spirit of Nagel and Singleton (2011).

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Cattaneo, M. D., Crump, R. K., Farrell, M. H., and Schaumburg, E. (2018). Characteristic-sorted portfolios: Estimation and inference. Staff Report 788, Federal Reserve Bank of New York.

Daniel, K., Mota, L., Rottke, S., and Santos, T. (2017). The cross-section of risk and return. Working Paper 24164, National Bureau of Economic Research.

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 322(10):891–921.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F., Ledoit, O., and Wolf, M. (2017). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*. doi: 0.1080/07350015.2017.1345683.

Engle, R. F. and Mezrich, J. (1996). GARCH for groups. *Risk*, 9:36–40.

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

Fama, E. F. and French, K. R. (2015). A five factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Green, J., Hand, J. R. M., and Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies*, 18:692–730.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Haugen, R. A. and Baker, N. L. (1991). The efficient market inefficiency of capitalization-weighted stock portfolios. *Journal of Portfolio Management*, 17(3):35–40.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705.

Korajczyk, R. A. and Sadka, R. (2004). Are momentum profits robust to trading costs? *The Journal of Finance*, 59(3):1039–1082.

Ledoit, O. and Péché, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields*, 150(1–2):233–264.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603–621.

Ledoit, O. and Wolf, M. (2004a). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30(4):110–119.

Ledoit, O. and Wolf, M. (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Annals of Statistics*, 40(2):1024–1060.

Ledoit, O. and Wolf, M. (2015). Spectrum estimation: a unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis*, 139(2):360–384.

Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Review of Financial Studies*, 30(12):4349–4388.

Lo, A. W. and MacKinlay, A. C. (1990). Data-snooping biases in tests of financial asset pricing models. *The Review of Financial Studies*, 3(3):431–467.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7:77–91.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *Journal of Finance*, 71(1):5–32.

Michaud, R. (1989). The Markowitz optimization enigma: Is optimized optimal? *Financial Analysts Journal*, 45:31–42.

Nagel, S. and Singleton, K. J. (2011). Estimation and evaluation of conditional asset pricing models. *Journal of Finance*, 66(3):873–909.

Pakel, C., Shephard, N., Sheppard, K., and Engle, R. F. (2017). Fitting vast dimensional time-varying covariance models. Working Paper FIN-08-009, NYU. Available at `https://ssrn.com/abstract=1354497`.

Scherer, B. (2011). A note on the returns from mininum variance investing. *Journal of Empirical Finance*, 18:652–660.

Silverstein, J. W. and Bai, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *Journal of Multivariate Analysis*, 54:175–192.

Stein, C. (1986). Lectures on the theory of estimation of many parameters. *Journal of Mathematical Sciences*, 34(1):1373–1403.

Welch, I. (2008). The link between Fama-French time-series tests and Fama-Macbeth cross-sectional tests. Technical report, UCLA Anderson School of Management. Available at https://ssrn.com/abstract=1271935.

A Tables and Figures

Table 1: List of factors

Number	Name
1	11-month momentum, 11-MM
2	1-month momentum (reversal), 1-MM
3	6-month momentum, 6-MM
4	Maximum daily return in prior month (reversal), Mxret
5	Change in 6-month momentum (reversal), Δ6-MM
6	Cumulative abnormal stock returns around earnings announcement, Abr
7	Dollar trading volume from month $t-2$ (reversal), Dvol
8	Log firm size (reversal), logME
9	Book-to-market, B/M
10	Asset growth, Agr
11	Earnings-to-price, E/P
12	Change in long-term debt (reversal), Δlgr
13	Change in common shareholder equity, Δceq
14	Cash flow from operation, Cflow
15	Cash-to-price (reversal), Cash
16	Dividend yield, D/P
17	Payout yield, O/P
18	Net payout yield, NO/P
19	Sales growth, SG
20	Market leverage, A/ME
21	Abnormal volume in earnings announcement month, Aevol
22	Earnings surprise, Sue
23	Change in order backlog, OB
24	Working capital accrual (reversal), Acc
25	Capital expenditures and inventory (reversal), Δcapx
26	Changes in inventory (reversal), Cii
27	Abnormal corporate investment (reversal), Aci
28	Net stock issues (reversal), Nsi
29	Net operating assets (reversal), Noa
30	Investment growth (reversal), IG
31	Net external financing (reversal), Nxf
32	Composite issuance (reversal), Cei
33	Total accruals (reversal), TA/A
34	Inventory growth (reversal), Ivg
35	Percent operating accruals (reversal), Poa

Table 1 continued

Number	Name
36	Percent total accruals (reversal), Pta
37	Change in deferred revenues, Δdrev
38	F-score
39	Change in profit margin, ΔPM
40	Asset turnover, Ato
41	Change in tax expense, Δtax
42	Return on assets, Roa
43	Gross profits-to-assets, Gma
44	Return on invested capital, Roic
45	Return on equity, Roe
46	Return on net operating assets, Rna
47	Taxable income-to-book income, TI/BI
48	Capital turnover, Cto
49	O-score
50	Operating profitability, OP
51	Employee growth rate (reversal), Egr
52	Change in advertising expense, Δade
53	R&D increase, Rdi
54	Advertisement expense-to-market, Ad/M
55	R&D-to-sales, RD/S
56	R&D-to- market, RD/M
57	R&D capital-to-assets, Rc/A
58	Operating leverage, OL
59	Turn (reversal)
60	Total Volatility (reversal), Tvol
61	Accrual Volatility (reversal), Avol
62	Cash flow volatility (reversal), Cvol

Table 2: t-statistics and their ratios. The columns labeled Qu contain the t-statistics (3.10) for quantile-based weighting; the columns labeled Ef contain the test statistics (3.10) for efficient weighting; the columns labeled Ef/Qu contain the corresponding ratios Ef/Qu for the cases when Qu is positive. NaN denotes missing values due to lack of sufficient data. NoI stands for "Not of Interest" and corresponds to cases when the t-statistic for Qu is negative.

Number	$N = 100$			$N = 500$			$N = 1000$		
	Qu	Ef	Ef/Qu	Qu	Ef	Ef/Qu	Qu	Ef	Ef/Qu
1	1.441	2.611	1.81	1.398	1.758	1.26	1.402	1.671	1.19
2	0.115	2.664	23.18	0.816	4.862	5.96	1.074	5.064	4.71
3	−0.296	0.051	NoI	0.312	-0.397	1.27	0.560	−0.595	1.06
4	0.483	0.710	1.47	−0.156	−0.820	NoI	−0.498	−2.812	NoI
5	0.217	0.623	2.88	1.198	2.422	2.02	1.342	2.433	1.81
6	1.259	1.698	1.35	2.658	3.066	1.15	3.219	4.460	1.39
7	−0.529	0.222	NoI	1.612	4.202	2.61	2.998	4.100	1.37
8	0.219	1.092	4.99	1.147	4.598	4.01	2.323	4.992	2.15
9	−0.123	−0.550	NoI	0.640	1.116	1.74	1.047	1.736	1.66
10	−0.056	0.249	NoI	0.016	0.235	14.62	0.573	1.050	1.83
11	2.411	4.672	1.94	4.544	11.085	2.44	5.345	15.716	2.94
12	0.517	1.854	3.58	0.849	2.875	3.39	2.056	4.257	2.07
13	1.006	0.717	0.71	2.671	4.075	1.53	3.308	7.862	2.38
14	5.306	6.088	1.15	6.713	9.825	1.46	7.108	16.031	2.26
15	1.820	3.361	1.85	2.807	5.667	2.02	3.864	6.434	1.67
16	−0.417	1.201	NoI	−0.291	0.995	NoI	−1.160	0.399	NoI
17	0.857	1.519	1.77	0.892	2.888	3.24	0.726	3.418	4.71
18	0.729	1.467	2.01	0.503	3.373	6.70	0.536	4.981	9.29
19	0.282	1.066	3.78	1.533	4.779	3.12	2.752	7.416	2.69
20	−0.661	−0.922	NoI	0.133	0.388	2.91	0.451	0.884	1.96
21	0.889	1.028	1.16	2.212	1.976	0.89	2.259	4.263	1.89
22	2.417	3.260	1.35	4.854	10.062	2.07	8.116	16.914	2.08
23	−0.064	−0.180	NoI	−0.300	1.683	NoI	NaN	NaN	NaN
24	3.046	4.703	1.54	5.102	7.006	1.37	7.363	10.803	1.47
25	0.631	1.883	2.98	1.579	3.964	2.51	3.287	5.050	1.54
26	1.340	2.221	1.66	1.886	3.748	1.99	2.715	4.589	1.69
27	1.406	2.581	1.84	3.346	3.975	1.19	3.760	5.354	1.42
28	−0.382	1.507	NoI	1.531	2.718	1.78	1.411	3.437	2.44
29	2.741	1.823	0.67	3.697	4.012	1.086	3.486	5.296	1.52
30	0.929	1.499	1.61	2.461	4.305	1.75	2.759	4.033	1.46
31	2.309	1.548	0.67	2.595	2.766	1.07	2.726	5.671	2.08
32	1.136	1.089	0.96	1.647	3.756	2.28	NaN	NaN	NaN

Table 2 continued

Number	$N = 100$			$N = 500$			$N = 1000$		
	Qu	Ef	Ef/Qu	Qu	Ef	Ef/Qu	Qu	Ef	Ef/Qu
33	0.020	1.048	52.40	1.857	3.354	1.81	3.068	3.422	1.12
34	1.516	2.138	1.41	1.874	4.336	2.31	2.945	4.801	1.63
35	1.736	2.975	1.71	2.174	3.461	1.59	4.229	6.919	1.64
36	1.397	1.711	1.22	1.555	3.418	2.20	2.450	3.249	1.33
37	2.257	1.069	0.47	3.491	4.098	1.17	NaN	NaN	NaN
38	0.541	1.304	2.41	1.505	3.097	2.06	1.368	4.478	3.27
39	2.012	2.500	1.24	3.482	7.704	2.21	5.778	11.764	2.04
40	1.427	2.452	1.72	2.339	3.259	1.39	2.802	4.576	1.63
41	1.761	2.924	1.66	4.557	8.957	1.97	6.968	15.678	2.25
42	2.302	3.641	1.58	3.538	6.453	1.82	4.459	10.142	2.27
43	1.963	3.798	1.93	2.424	4.320	1.78	2.964	6.265	2.11
44	2.435	4.010	1.65	3.105	5.941	1.91	4.165	9.310	2.24
45	2.537	3.207	1.26	4.297	7.340	1.71	4.975	11.897	2.39
46	3.243	4.532	1.40	3.869	5.956	1.54	4.506	9.812	2.18
47	1.414	2.424	1.71	1.031	2.626	2.55	0.752	3.208	4.26
48	1.822	1.964	1.08	1.605	2.543	1.58	2.435	3.876	1.59
49	−2.158	−0.915	NoI	−1.474	−0.620	NoI	0.532	−1.696	−3.19
50	2.330	2.543	1.09	3.562	4.340	1.22	4.353	9.080	2.09
51	0.573	0.350	0.61	0.802	2.028	2.53	1.261	2.701	2.14
52	0.919	0.320	0.35	−0.011	0.657	NoI	NaN	NaN	NaN
53	−0.638	0.174	NoI	−0.614	−0.320	NoI	NaN	NaN	NaN
54	0.213	1.200	5.63	2.018	1.110	0.55	NaN	NaN	NaN
55	0.719	1.204	1.67	1.550	3.333	2.15	NaN	NaN	NaN
56	1.553	1.312	0.84	3.348	4.737	1.42	NaN	NaN	NaN
57	1.132	1.762	1.56	1.960	5.521	2.82	NaN	NaN	NaN
58	1.463	1.899	1.30	1.718	2.370	1.38	2.675	3.376	1.26
59	−0.211	−0.882	NoI	−0.347	0.213	NoI	−0.175	0.691	NoI
60	0.114	1.244	10.95	−0.251	0.408	NoI	−0.548	−1.158	NoI
61	1.653	1.172	0.71	1.138	1.087	0.95	0.602	1.955	3.25
62	2.486	1.738	0.70	2.758	2.257	0.82	2.862	2.999	1.05

N	Qu > 0	Qu > 0.5	Qu > 1.0
100	3.32	1.45	1.34
500	2.23	2.04	1.79
1000	2.06	2.07	1.95

Table 3: Averages based on the columns labeled Ef/Qu in Table 2. The second column reports averages when the t-statistic for Qu is positive; the third column reports averages when the t-statistic for Qu is greater than 0.5; and the fourth column reports averages when the t-statistic for Qu is greater than 1.0.

Critical value = 2				
N	Qu	Ef	Qu	Ef
100	14	22	0.23	0.35
500	25	46	0.40	0.74
1000	34	42	0.64	0.79

Critical value = 3				
N	Qu	Ef	Qu	Ef
100	3	10	0.05	0.16
500	15	36	0.24	0.58
1000	19	39	0.36	0.74

Table 4: Number (columns two and three) and proportion (columns four and five) of the t-statistics in Table 2 whose value exceed two (left panel) and three (right panel), respectively.

N	Qu > 0	Qu > 0.5	Qu > 1.0
100	3.55	0.98	0.88
500	3.89	1.67	1.52
1000	3.44	1.61	1.66

Table 5: Similar to Table 3, except that these results are for the portfolio returns filtered through the Fama-French five-factor model.

Critical value = 2					Critical value = 3				
N	Qu	Ef	Qu	Ef	N	Qu	Ef	Qu	Ef
100	22	22	0.35	0.35	100	11	8	0.18	0.13
500	30	44	0.48	0.71	500	25	34	0.40	0.55
1000	30	42	0.57	0.79	1000	25	38	0.47	0.72

Table 6: Similar to Table 4, except that these results are for the portfolio returns filtered through the Fama-French five-factor model.

	FF Factors	Ef Factors
$\hat{\gamma}_0^*$	0.08	0.08
t	3.58	3.60
pv	0.00	0.00

Table 7: This table extends the analysis of Welch (2008, Section IV) from 6 size and value portfolios to 25. The first row ($\hat{\gamma}_0^*$) contains the point estimates for the parameter γ_0^* in his regression model (8); the second row (t) contains the corresponding t-statistics; and the third row (pv) contains the bootstrap p-values for the null hypothesis $H_0 : \gamma_0^* = 0$. The first column (FF Factors) contains the results for the three (demeaned) original Fama-French factors as regressors in his model (8); the second column (Ef Factors) contains the results when our (demeaned) 'efficient' factors 8 and 9 are used as size and value regressors, respectively, instead.

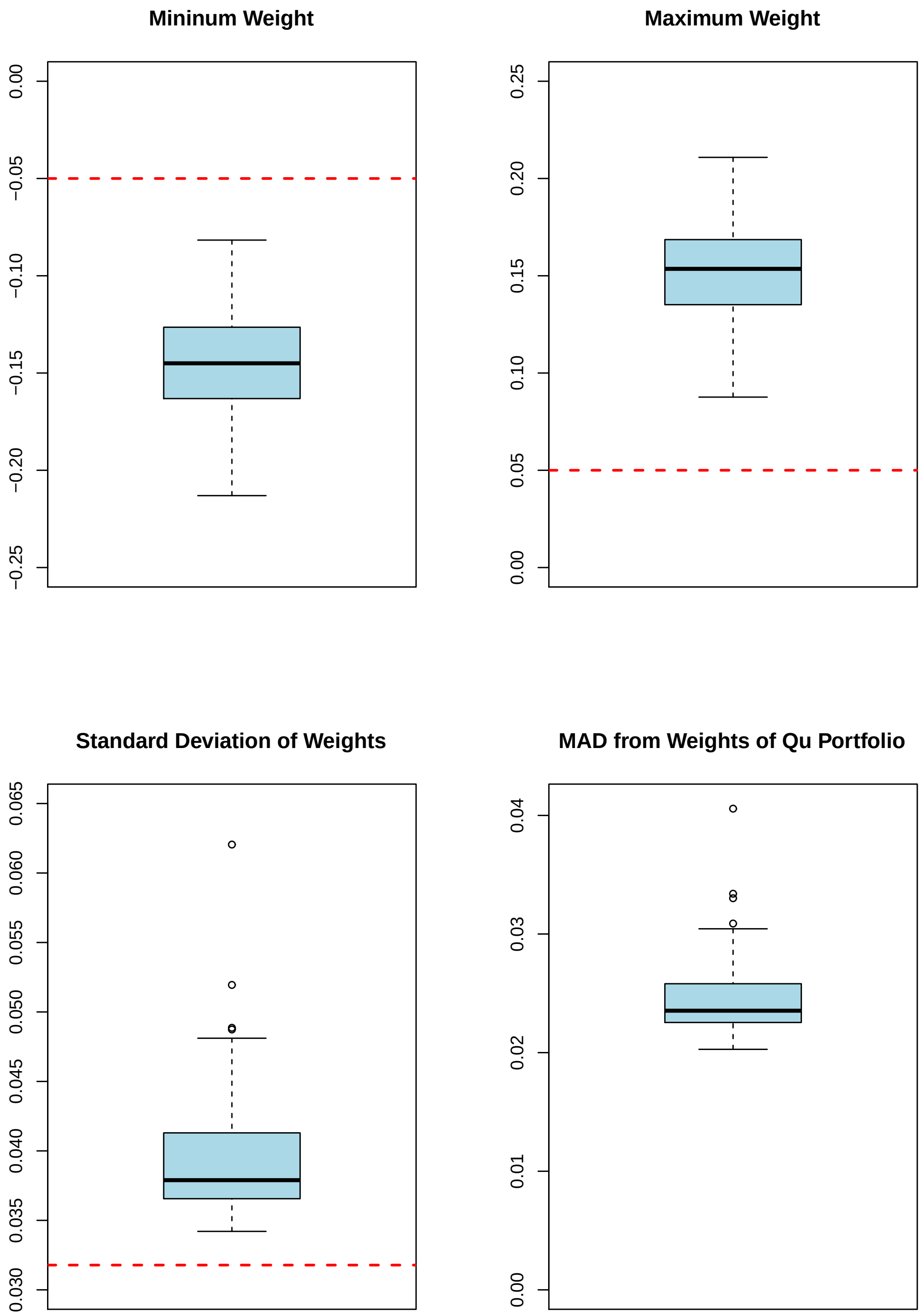

Figure 1: Boxplots of summary statistics of the Ef portfolio weights for the 62 factors under consideration for $N = 100$. (In the first three panles, the dashed line indicates the corresponding, constant summary statistic for the Qu portfolio weights.)

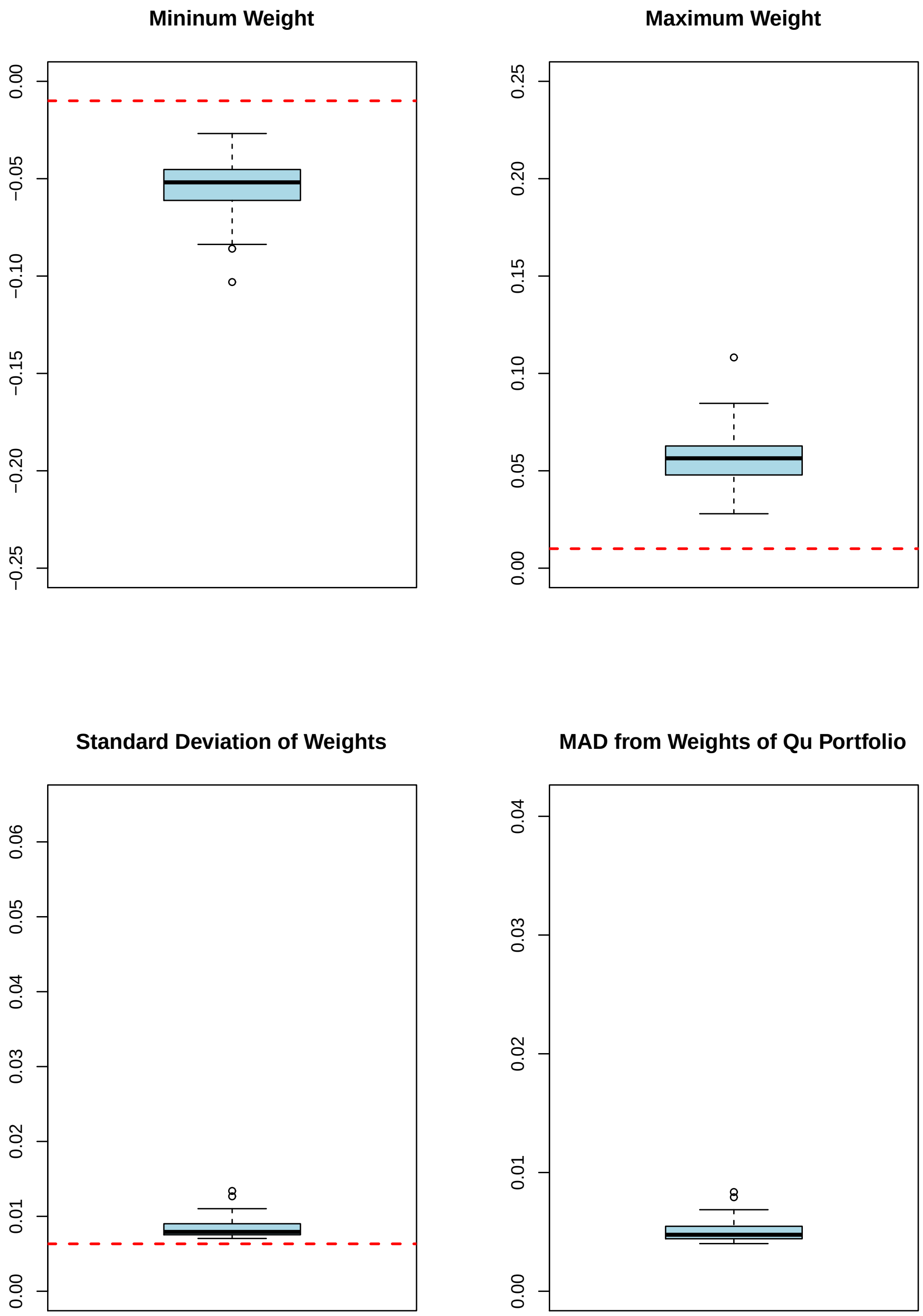

Figure 2: Boxplots of summary statistics of the Ef portfolio weights for the 62 factors under consideration for $N = 500$. (In the first three panles, the dashed line indicates the corresponding, constant summary statistic for the Qu portfolio weights.)

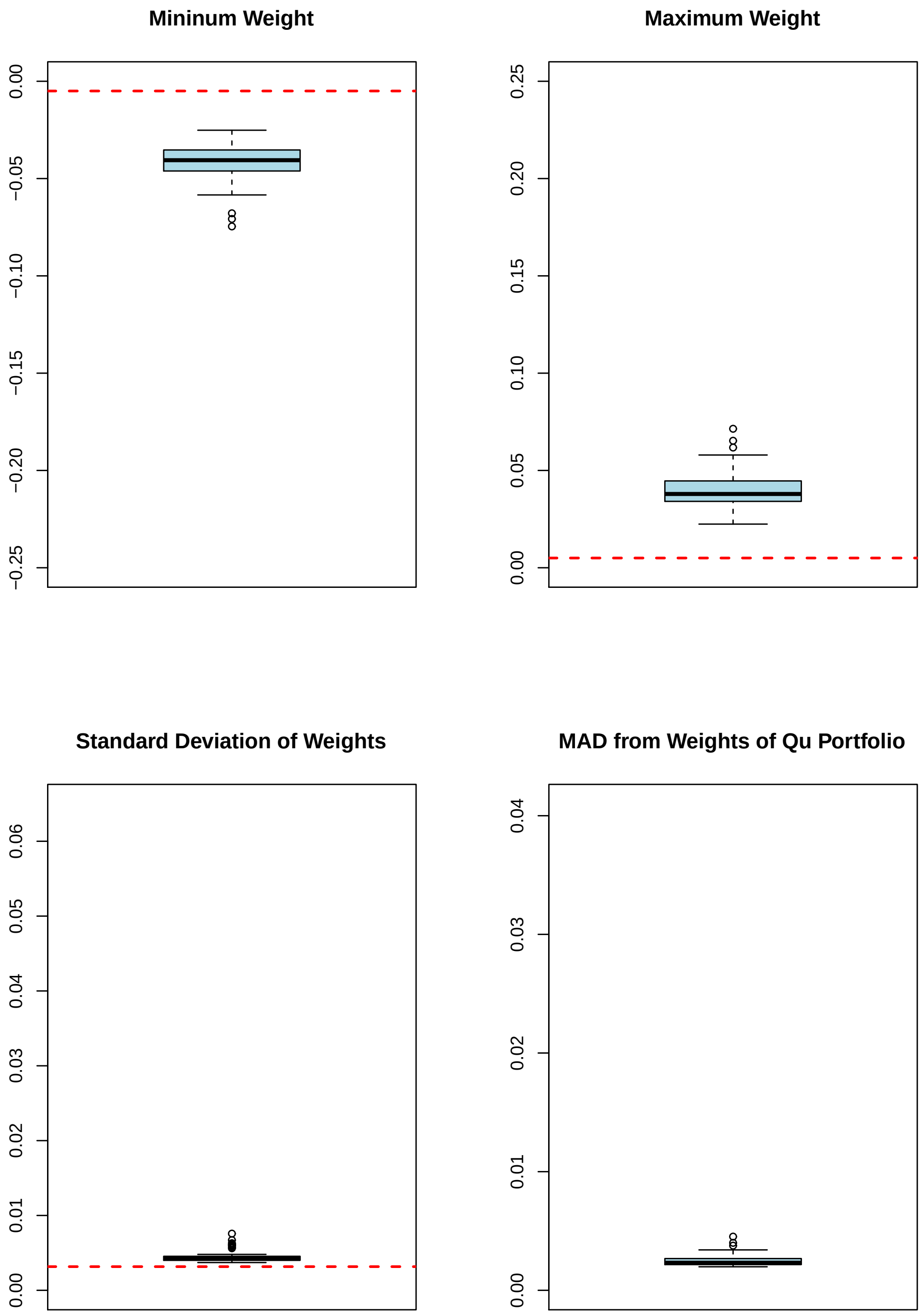

Figure 3: Boxplots of summary statistics of the Ef portfolio weights for the 53 factors under consideration for $N = 1000$. (In the first three panles, the dashed line indicates the corresponding, constant summary statistic for the Qu portfolio weights.)

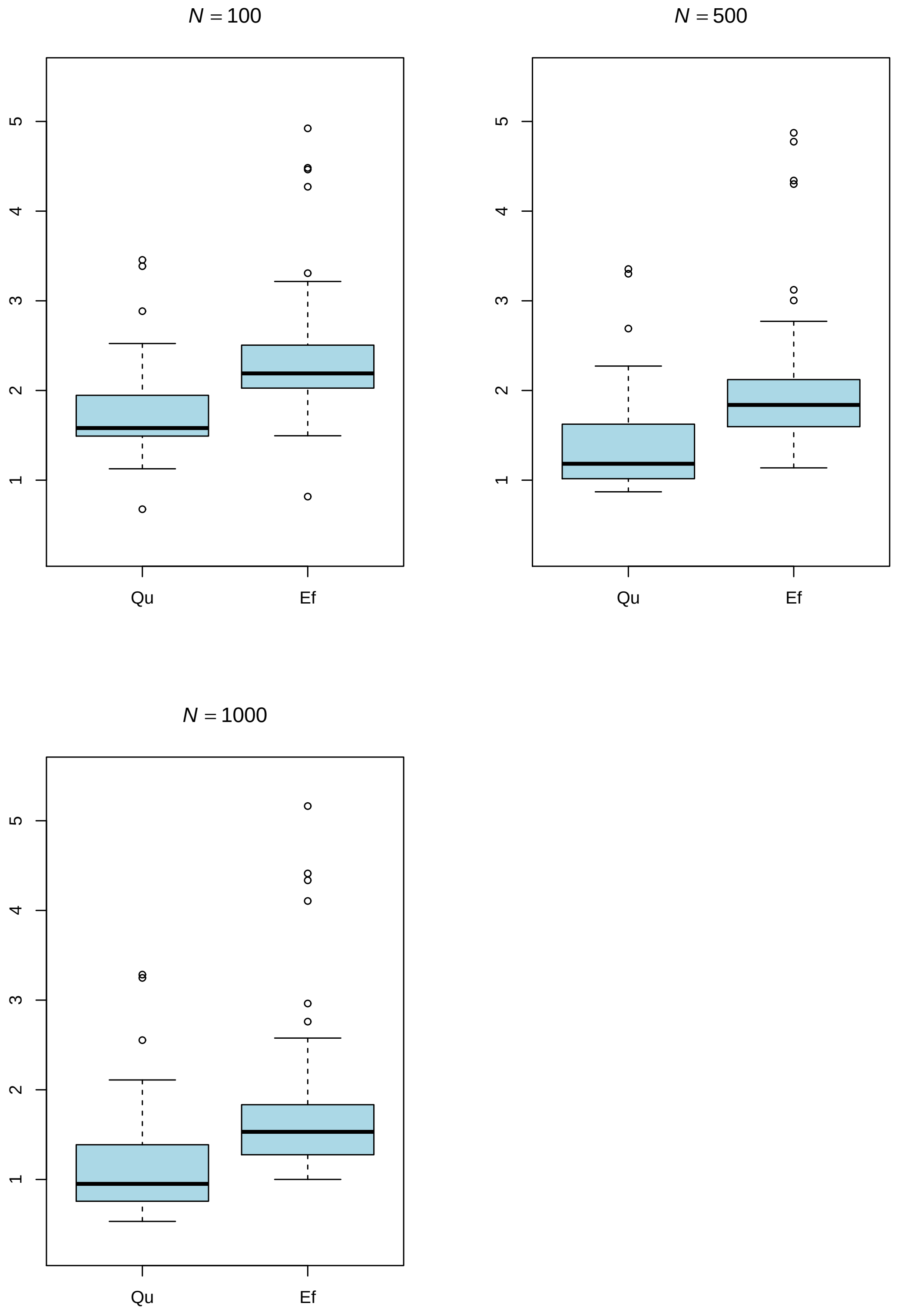

Figure 4: Boxplots of average turnover for the factors under consideration. There are 62 factors for $N \in \{100, 500\}$ and 53 factors for $N = 1000$.

B Winsorization of Factor Scores

'Outlying' factor scores that are unusually large in magnitude can have undesirable impacts when used as input in Markowitz optimization. We migitate this potential problem by properly truncating very small and very large values in any cross-sectional vector of factor scores $\boldsymbol{m}_t$. Such truncation is commonly referred to as 'Winsorization', a method that is widely used by quantitative portfolio managers; for example, see (Chincarini and Kim, 2006, Appendix 5B).

Consider a generic vector $\boldsymbol{a} := (a_1, \ldots, a_N)'$. We first compute a robust measure of location that is not (heavily) affected by potential outliers. To this end, we use the trimmed mean of the data with trimming fraction $\eta \in (0, 0.5)$ on the left and on the right. This number is simply the mean of the middle $(1 - 2\eta) \cdot 100\%$ of the data. More specifically, denote by

$$a_{(1)} \leq a_{(2)} \leq \ldots \leq a_{(N)} \tag{B.1}$$

the ordered data (from smallest to largest) and denote by

$$M := \lfloor \eta \cdot N \rfloor \tag{B.2}$$

the smallest integer less than or equal to $\eta \cdot N$. Then the trimmed mean with trimming fraction η is defined as

$$\overline{\boldsymbol{a}}_\eta := \frac{1}{N - 2M} \sum_{i=M+1}^{N-M} a_{(i)} \ . \tag{B.3}$$

We employ the value of $\eta = 0.1$ in practice.

We next compute a robust measure of spread. To this end, we use the median absolute deviation (MAD) from the median given by

$$\text{MAD}(\boldsymbol{a}) := \text{med}\big(|\boldsymbol{a} - \text{med}(\boldsymbol{a})|\big) \ . \tag{B.4}$$

where $\text{med}(\cdot)$ denotes the sample median of a vector and $|\cdot|$ denotes the element-wise absolute-value function of a vector.

We next compute upper and lower bounds defined by

$$a_{\text{lo}} := \overline{\boldsymbol{a}}_{0.1} - 5 \cdot \text{MAD}(\boldsymbol{a}) \quad \text{and} \quad a_{\text{up}} := \overline{\boldsymbol{a}}_{0.1} + 5 \cdot \text{MAD}(\boldsymbol{a}) \ . \tag{B.5}$$

The motivation here is that for a normally distributed sample, it will hold that $\overline{a} \approx \overline{a}_{0.1}$ and $s(\boldsymbol{a}) \approx 1.5 \cdot \text{MAD}(\boldsymbol{a})$, where $\overline{\boldsymbol{a}}$ and $s(\boldsymbol{a})$ denote the sample mean and the sample median of $a_1, \ldots, a_N$, respectively. As a result, for a 'well-behaved' sample, there will usually be no points below a_{lo} or above a_{up}. Our final truncation rule is that any data point a_i below a_{lo} will be changed to a_{lo} and any data point a_i above a_{up} will be changed to a_{up}.

We then apply this truncation rule to the cross-sectional vector of factor scores $\boldsymbol{m}_t$ in place of the generic vector $\boldsymbol{a}$.

C Description of Factors

Daily data used are from the Center for Research in Security Prices (CRSP), including holding period returns (item ret), return without dividends (item retx), prices (item prc), number of shares traded (item vol), number of shares outstanding (item csho), factor to adjust shares (item ajex), and value-weighted return (item vwretd). The other data are from the Compustat Annual and Quarterly Fundamental Files. For each factor, we describe how the factor scores are computed at a generic investment date $h = 1, \ldots, 360$.

C.1 Momentum

C.1.1 11-MM

Following Fama and French (1996), we calculate 11-month momentum (11-MM) as the average return over the previous 12 months but excluding the most recent month. That is, we compute the average return from day $h - 252$ through day $h - 22$.

C.1.2 1-MM

Following Jegadeesh and Titman (1993), we calculate 1-month momentum (1-MM) as the average return from day $h - 21$ through day $h - 1$. Reversal of 1-MM (that is, the negative of 1-MM) is used as the actual factor.

C.1.3 6-MM

Following Jegadeesh and Titman (1993), we calculate 6-month momentum (6-MM) as the average return over the previous seven months but excluding the most recent month. That is, for any investment date date h, we compute the average return from day $h - 147$ through day $h - 22$.

C.1.4 Mxret

Following Bali et al. (2011), Mxret is the maximum daily return from day $h - 21$ through day $h - 1$. Reversal of Mxret is used as the actual factor.

C.1.5 Δ6-MM

Following Gettleman and Marks (2006), change in 6 month momentum(Δ6-MM) is calculated as current 6-MM minus previous 6-MM (that is, 6-MM at investment date $h - 1$). Reversal of Δ6-MM is used as the actual factor.

C.1.6 Abr

Following Chan et al. (1996), we measure cumulative abnormal stock return (Abr) around the latest quarterly earnings announcement date as

$$\text{Abr}_i := \sum_{d=-2}^{1} (r_{id} - r_{md}) \ , \tag{C.1}$$

where r_{id} and r_{md} are, respectively, the return of stock i and the value-weighted return of the market index (item vwretd) on day d, where $d = 0$ represents the earnings announcement day (quarterly item rdq). For stock i, at every investment date h, we use the most recent earnings announcement day as long as the day is at least two days earlier than the investment day (to make sure that $r_{i(d=1)}$ is available).

C.2 Value-versus-Growth

C.2.1 Dvol

Dvol is the dollar trading volume in the latest-but-one month (that is, from day $h-42$ through day $h-22$). As in Chordia et al. (2001), we measure it as the natural log of the sum of daily dollar trading volume during that period. Daily dollar trading volume is share price (item prc) times the number of shares traded (item vol). Reversal of Dvol is used as the actual factor.

C.2.2 logME

Banz (1981) proposes firm size as a factor. We use the logarithm of market capitalization (ME) of one day before the investment day (that is, on day $h-1$) as firm size. ME is calculated as price (item prc) times shares outstanding (item csho). Reversal of logME is used as the actual factor

C.2.3 B/M

Rosenberg et al. (1985) propose book-to-market as a factor. We measure it as the ratio of book equity to market capitalization on the day before the investment day (that is, on day $h-1$); here, book equity is computed from the most recently announced quarterly data. Our measure of the book equity is the quarterly version of the annual book equity measure in Davis et al. (2000). In particular, it is the book value of common equity (item ceqq) plus the par value of preferred stock (item pstkq), plus balance-sheet deferred taxes and investment tax credit (item txditcq), and then minus the book value of preferred stock. We use redemption value (item pstkrq, zero if missing) for the book value of preferred stock.

C.2.4 Agr

To construct the Cooper et al. (2008) asset growth (Agr) factor, we divide the total assets (item atq) by 1-quarter-lagged total assets; item atq uses the most recently announced quarterly data. Reversal of Agr is used as the actual factor.

C.2.5 E/P

Following Basu (1983), earnings-to-price (E/P) is calculated as income before extraordinary items (item ibq) divided by the market capitalization (ME) on day $h-1$; item ibq uses the most recently announced quarterly data.

C.2.6 Δlgr

Following Scott et al. (2005), we measure change in long-term debt (Δlgr) as long-term debt (item lt) divided by 1-year-lagged long-term debt minus one; item lt uses the most recently announced quarterly data. Reversal of Δlgr is used as the actual factor.

C.2.7 Δceq

Following Scott et al. (2005), we measure change in common shareholder equity (Δceq) as common shareholder equity (item ceqq) divided by 1-quarter-lagged common shareholder equity minus one; item ceqq uses the most recently announced quarterly data.

C.2.8 Cflow

Following Houge and Loughran (2000), we define cash flow from operation (Cflow) as net cash flow from operations in the most recently announced quarter scaled by the average of total assets (item atq) for the two previous quarters. Instead of using the item oancf (net cash flow from operations) directly, we use net income (item niq) minus operating accruals (OA) because these items have a broader coverage than oancf, and they have quarterly data. To measure OA, we use the balance-sheet approach of Sloan (1996), that is,

$$\mathrm{OA} := (\Delta\mathrm{actq} - \Delta\mathrm{cheq}) - (\Delta\mathrm{lctq} - \Delta\mathrm{dlcq} - \Delta\mathrm{txpq}) - \mathrm{dpq} \,, \tag{C.2}$$

where Δ represents the change in the corresponding item, and items actq, cheq, lctq, dlcq, txpq, dpq are corresponding to the quarterly data of current assets, cash and cash equivalents, current liabilities, debt included in current liabilities (zero if missing), income taxes payable (zero if missing), depreciation and amortization(zero if missing), respectively. Note that the number of stocks for which this factor is available during the first eight investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 9$.

C.2.9 Cash

Following Chandrashekar and Rao (2009), cash to price (Cash) is computed as

$$\mathrm{Cash} := (\mathrm{ME} + \mathrm{dlttq} - \mathrm{atq})/\mathrm{cheq} \,, \tag{C.3}$$

where ME is the market capitalization on day $h-1$, and items dlttq, atq, and cheq are all quarterly data corresponding to long-term debt, total asset, and cash or cash equivalents, respectively; all these items use the most recently announced quarterly data. Reversal of Cash is used as the actual factor.

C.2.10 D/P

As in Litzenberger and Ramaswamy (1979), dividend yield (D/P) is measured as the total dividends paid out from the previous year (that is, from day $h-252$ through day $h-1$) divided by ME on day $h-1$. The total dividends are calculated by accumulating daily dividends, and the daily dividends is measured as the difference between cum- and ex-dividend returns, which are respectively corresponding to holding period returns (item ret) and return without dividends (item retx), times the 1-day-lagged ME.

C.2.11 O/P

Following Boudoukh et al. (2007), total payouts (O/P) are dividends on common stock (dvc) plus repurchases of the previous year (that is, from day $h-252$ through day $h-1$) divided by ME on day $h-1$. Repurchases are the total expenditure on the purchase of common and preferred stocks (item prstkc) minus the change over the previous year in the value of the net number of preferred stocks outstanding (item pstkrv).

C.2.12 NO/P

Following Boudoukh et al. (2007), net payouts (NO/P) are the same as total payouts except that the equity issuances have to be subtracted from the total payouts. Equity issuances are the sale of common and preferred stock (item sstk) minus the change over the previous year in the value of the net number of preferred stocks outstanding (item pstkrv).

C.2.13 SG

Lakonishok et al. (1994) propose sales growth (SG) as a factor. We measure it as the growth rate in sales (item saleq) from quarter $t-2$ through quarter $t-1$, where t denotes the current quarter.

C.2.14 A/ME

Following Bhandari (1988), A/ME is measured as the ratio of total assets in quarter $t-1$ to ME on day $h-1$, where t denotes the current quarter.

C.2.15 Aevol

As in Lerman et al. (2008), the abnormal earnings announcement period volume (Aevol) is defined as average daily share trading volume over the three days from $d=-1$ through $d=1$ divided by the average daily share volume over days $d=-8$ through $d=-63$, and then subtracting one, where $d=0$ denotes day of the most recent earnings announcement (item rdq):

$$\text{Aevol}_i := \frac{Avg_{d\in[-1,1]}(\text{vol}_{id})}{Avg_{d\in[-63,-8]}(\text{vol}_{id})} - 1 \ . \tag{C.4}$$

Note that the day of the most recent earnings announcement most be at least two days before the investment day h (to make sure that $\text{vol}_{i(d=1)}$ is available).

C.2.16 Sue

Following Foster et al. (1984), we measure earnings surprise (Sue) as the change in the most recently announced quarterly earnings per share (item epspxq) from its value four quarters ago, divided by the standard deviation of this change in quarterly earnings over the previous eight quarters.

C.2.17 OB

Following Gu et al. (2009), we measure OB as annual order backlog (item ob) in year $t-1$ scaled by the average of total assets (item at) for calendar years $t-2$ and $t-1$, where t denotes the current calendar year. Note that the number of stocks for which this factor is available during the first 65 investment periods is less than 500, and the number is less than 1000 for the entire investment period. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 66$ whereas for dimension $N = 1000$, we do not consider this factor.

C.3 Investment

Considering the general negative relation between investment and expected return, all factors in this section are used in reversal.

C.3.1 Acc

Following Sloan (1996), we measure working capital accruals (Acc) as operating accruals (OA) in quarter $t-1$ scaled by the average of total assets (item atq) for quarters $t-2$ and $t-1$, where t denotes the current quarter and OA is the same as in equation (C.2). Note that the number of stocks for which this factor is available during the first eight investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 9$.

C.3.2 Δcapx

Following Lyandres et al. (2008), we measure capital expenditures and inventory (Δcapx) as changes in gross property, plant, and equipment (item ppegt) plus changes in inventory (item invt) scaled by 1-year-lagged total assets (item at). Note that the number of stocks for which this factor is available during the first two investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 3$.

C.3.3 Cii

Following Thomas and Zhang (2002), we measure change in inventory (Cii) as the change in the most recently announced annual inventory from its value one year previous to that, scaled by the average of total assets (item at).

C.3.4 Aci

Following Titman et al. (2004), we measure abnormal corporate investment (Aci) as

$$\mathrm{Aci}_t := \frac{3*\mathrm{CE}_{t-1}}{\mathrm{CE}_{t-2}+\mathrm{CE}_{t-3}+\mathrm{CE}_{t-4}} - 1 \ , \tag{C.5}$$

where t denotes the current calendar year and CE_{t-j} is capital expenditure (item capx) scaled by sales (item sale) in calendar year $t-j$. Note that the number of stocks for which this factor is available during the first three investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 4$.

C.3.5 Nsi

Pontiff and Woodgate (2008) propose net stock issues (Nsi) as a factor. We measure it as the natural log of the ratio of the average split-adjusted shares outstanding over the previous year (that is, from day $h-252$ through day $h-1$) to the average split-adjusted shares outstanding over the year previous to that (that is, from day $h-504$ through day $h-253$). We measure the daily split-adjusted shares outstanding as shares outstanding (item csho) times the adjustment factor (item ajex).

C.3.6 Noa

As in Hirshleifer et al. (2004), we measure net operating assets (Noa) as operating assets minus operating liabilities. Operating assets are total assets (item atq) minus cash and short-term investment (item cheq). Operating liabilities are total assets minus debt included in current liabilities (item dlcq, zero if missing), minus long-term debt (item dlttq, zero if missing), minus minority interests (item mibq, zero if missing), minus preferred stocks (item pstkq, zero if missing), and minus common equity (item ceqq). We use quarterly data instead of annual data.

C.3.7 IG

Following Xing (2008), we measure investment growth (IG) as the growth rate in capital expenditure (item capx) from calendar year $t-2$ to calendar year $t-1$, where t denotes the current calendar year.

C.3.8 Nxf

Following Bradshaw et al. (2006), we measure net external financing (Nxf) as the sum of net equity financing and net debt financing in year calendar $t-1$ scaled by the average of total assets, where t denotes the current calendar year. Net equity financing is the proceeds from the sale of common and preferred stocks (item sstk) less cash payments for the repurchases of common and preferred stocks (item prstkc) less cash payments for dividends (item dv).

Net debt financing is the cash proceeds from the issuance of long-term debt (item dltis) less cash payments for long-term debt reductions (item dltr) plus the net changes in current debt (item dlcch, zero if missing). Note that the number of stocks for which this factor is available during the first 13 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 14$.

C.3.9 Cei

Following Daniel and Titman (2006), we define composite issuance (Cei) as the growth rate in market capitalization (ME) during the previous five years (that is, from day $h-1260$ through day $h-1$) not attributable to the stock return. It is calculated as

$$\mathrm{Cei}_t := \log(\mathrm{ME}_t - \mathrm{ME}_{t-5}) - \log r(t-5, t) \ , \tag{C.6}$$

where $r(t-5,t)$ is the cumulative return on the stock from day $h-1260$ through day $h-1$, ME_t is the ME on day $h-1$, and ME_{t-5} is the ME on day $h-1260$. Note that the number of stocks for which this factor is available during some middle investment periods (for example, from 08/29/2011 through 12/31/2012) is less than 1000. As a result, for dimension $N = 1000$, we do not consider this factor.

C.3.10 TA/A

Following Richardson et al. (2005), we measure TA/A as total accruals scaled by 1-year-lagged total assets (item at). Total accruals (TA) are calculated as

$$\mathrm{TA} := \varDelta\mathrm{WC} + \varDelta\mathrm{NCO} + \varDelta\mathrm{FIN} \ , \tag{C.7}$$

where $\varDelta$ represents the change in the corresponding item, and items WC, NCP, FIN are net non-cash working capital, net non-current operating assets, and net financial assets, respectively:

$$\mathrm{WC} := \mathrm{act} - \mathrm{che} - (\mathrm{lct} - \mathrm{dlc}) \tag{C.8}$$
$$\mathrm{NCO} := \mathrm{at} - \mathrm{act} - \mathrm{ivao} - (\mathrm{lt} - \mathrm{lct} - \mathrm{dltt}) \tag{C.9}$$
$$\mathrm{FIN} := \mathrm{ivst} + \mathrm{ivao} - (\mathrm{dltt} + \mathrm{dlc} + \mathrm{pstk}) \ . \tag{C.10}$$

Here, act, che, lct, dlc, at, ivao, lt, lct, dltt, ivst, pstk are all annual items corresponding to current assets, cash and short-term investment, current liabilities, debt in current liabilities, total assets, long-term investments (zero if missing), total liabilities, current liabilities, long-term debt (zero if missing), short-term investment (zero if missing), and preferred stock (zero if missing), respectively. Note that the number of stocks for which this factor is available during the first 5 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 6$.

C.3.11 Ivg

Following Belo and Lin (2012), we define inventory growth (Ivg) as the growth rate of inventory (item invt) from calendar year $t-2$ to year calendar year $t-1$, where t denotes the current calendar year.

C.3.12 Poa

Following Hafzalla et al. (2011), percent operating accruals (Poa) is measured as operating accruals (OA) in quarter $t-1$, scaled by net income (item niq) in the same quarter, where t denotes the current quarter; see equation (C.2) for the definition of OA. Note that the number of stocks for which this factor is available during the first eight investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 9$.

C.3.13 Pta

Following Hafzalla et al. (2011), percent total accruals (Pta) is measured as total accruals (TA) scaled by net income (item ni); see equation (C.7) for the definition of TA. Considering the broader coverage, we use annual data instead of quarterly data to calculate this factor. Note that the number of stocks for which this factor is available during the first 6 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 7$.

C.4 Profitability

C.4.1 Δdrev

Following Prakash and Sinha (2013), we measure change in deferred revenues (Δdrev) as the growth rate of deferred revenues (item drcq) from quarter $t-2$ to quarter $t-1$, where t denotes the current quarter. Note that the number of stocks for which this factor is available is less than 1000 during the entire investment period; therefore, we do not consider this factor for dimension $N = 1000$. According to the available number of stocks, for dimension $N = 100$, we start the portfolio formation on investment date $h = 221$ whereas for dimension $N = 500$, we start the portfolio formation on investment date $h = 229$.

C.4.2 F-score

Following Piotroski (2000), we define F-score as the sum of nine individual binary signals:

$$F := F_{Roa} + F_{\Delta Roa} + F_{Cfo} + F_{Acc} + F_{\Delta Margin} + F_{\Delta Turn} + F_{\Delta Lever} + F_{\Delta Liquid} + F_{EQ} \quad \text{(C.11)}$$

where Roa is income before extraordinary (item ib) scaled by 1-year-lagged total assets (item at); ΔRoa is the increase in Roa compared to the previous year; Cfo is cash flow from operation (we use funds from operation (item fopt) minus the annual change in working capital (item wcap) scaled by 1-year-lagged total assets; Acc is defined as Cfo minus Roa; $\Delta Margin$ is gross margin (item sale minus cogs, and then divided by sale) in calendar year $t-1$ less gross margin in calendar year $t-2$; $\Delta Turn$ is the change in the current calendar year's asset turnover ratio, which is measured as total sales (item sale) scaled by 1-year-lagged total assets (item at), compared to the previous calendar year; $\Delta Lever$ is the decrease in the current calendar year's lever, which is measured as total long-term debt (item dltt) scaled by average total assets over the previous two calendar years; $\Delta Liquid$ is the change in the current calendar year's current ratio compared to the previous calendar year, which is measured as the ratio of current assets (item act) to current liabilities (item lct); EQ, which measures whether the firm issue common equity in the current calendar year, equals the increase in preferred stock (item pstk) minus the sales of common and preferred stocks (item sstk). For our definition, the indicator variable always is equal to 1 if the corresponding variable is positive and is equal to zero otherwise.

C.4.3 ΔPM

Following Soliman (2008), we measure change in profit margin (ΔPM) as profit margin in quarter $t-1$ less profit margin in quarter $t-2$, where t denotes the current quarter. Profit margin is operating income after depreciation (item oiadp), scaled by sales (item saleq).

C.4.4 Ato

Following Soliman (2008), we measure asset turnover (Ato) as sales (quarterly item saleq), divided by 1-quarter-lagged Noa (net operating assets); see Section C.3.6 for a description of Noa.

C.4.5 Δtax

Following Thomas and Zhang (2011), we measure changes in tax expense (Δtax) as tax expense (item txtq) in quarter t minus tax expense in quarter $t-4$, scaled by total assets (item atq) in quarter $t-4$, where t denotes the current quarter.

C.4.6 Roa

Following Balakrishnan et al. (2010), we measure return on assets (Roa) as income before extraordinary items (item ibq) divided by 1-quarter-lagged total assets (item atq).

C.4.7 Gma

Following Novy-Marx (2010), we measure Gross profitability (Gma) as sales (item saleq) minus cost of goods sold (item cogsq), then divided by 1-quarter-lagged total assets (item atq).

C.4.8 Roic

Following Brown and Rowe (2007), we measure return on invested capital (Roic) as operating income after depreciation (quarterly item oiadpq) divided by 1-quarter-lagged operating assets, which are total assets (item atq) minus cash and short-term investment (item cheq).

C.4.9 Roe

Following Haugen and Baker (1996), we measure return on equity (Roe) as income before extraordinary items (quarterly item ibq) divided by 1-quarter-lagged book equity; book equity is computed as in Section C.2.3.

C.4.10 Rna

Following Soliman (2008), we measure return on operating assets (Rna) as operating income after depreciation (quarterly item oiadpq) divided by 1-quarter-lagged net operating assets (Noa); see Section C.3.6 for a description of Noa.

C.4.11 TI/BI

Following Green et al. (2014), we measure taxable income-to-book income (TI/BI) as pretax income (quarterly item piq) divided by net income (item niq).

C.4.12 Cto

Following Haugen and Baker (1996), we measure capital turnover (Cto) as sales (quarterly item saleq) divided by 1-quarter lagged total assets (item atq).

C.4.13 O-score

Following Ohlson (1980), we define the O-score as

$$\begin{aligned} \mathrm{O} := & -1.32 - 0.407\log(\mathrm{at}) + 6.03\mathrm{tlta} - 1.43\mathrm{wcta} + 0.076\mathrm{clca} \\ & -1.72\mathrm{oeneg} - 2.37\mathrm{nita} - 1.83\mathrm{futl} + 0.285\mathrm{intwo} - 0.521\mathrm{chin} \end{aligned} \tag{C.12}$$

where tlta := (dlc+dltt)/at, wcta := (act-lct)/at, clca:=lct/act, nita:=ni/at, and futl:=pi/lt. that oeneg is equal to 1 if lt exceeds at and is equal to zero otherwise. intwo is equal to 1 if ni for the last two calendar years is negative and is equal to zero otherwise. chine $= (\mathrm{ni}_t - \mathrm{ni}_{t-1})/(|\mathrm{ni}_t| + |\mathrm{ni}_{t-1}|)$. at, dlc, dltt, act, lct, ni, pi, lt are all annual items corresponding to total assets, debt in current liabilities, long-term debt, current assets, current liabilities, net income, pretax income, and total liabilities, respectively. Note that the number of stocks for which this factor is available during the first 5 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 6$.

C.4.14 OP

Following Fama and French (2015), we measure operating profitability (OP) with accounting data for quarter $t-1$ as revenues (item saleq) minus cost of goods sold (item cogsq), minus selling, general, and administrative expenses (item (item xsgaq), minus interest expense (item xintq) all divided by book equity. Book equity is the same as described in Section C.2.3.

C.5 Intangibles

C.5.1 Egr

Following Bazdrech et al. (2008), we measure employee growth rate (Egr) as the growth rate in the number of employees (item emp) from calendar year $t-2$ to calendar year $t-1$, where t denotes the current calendar year. Reversal of Egr is used as the actual factor.

C.5.2 Δade

Following Chemmanur and Yan (2010), we measure change in advertising expense (Δade) as the the natural log of the ratio of advertising expenses in calendar year $t-1$ to advertising expenses in calendar year $t-2$, where t denotes the current calendar year. Note that the number of stocks for which this factor is available during some of the first 181 investment periods is less than 500, and the number available from the 182th investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 182$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.3 Rdi

Following Eberhart et al. (2004), we measure R&D increase (Rdi) as the growth rate in R&D expenses (item xrd) from calendar year $t-2$ to calendar year $t-1$, where t denotes the current calendar year. Note that the number of stocks for which this factor is available during some of the first 26 investment periods is less than 500, and the number available from the 27th investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 27$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.4 Ad/M

As in Chan et al. (2001), we measure advertisement expense-to-market (Ad/M) as advertising expenses (item xad) for calendar year $t-1$ divided by the market capitalization (ME) on day $h-1$, where t denotes the current calendar year. Note that the number of stocks for which this factor is available during some of the first 169 investment periods is less than 500, and the number available from the 170th investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 170$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.5 RD/S

Following Chan et al. (2001), we measure R&D-to-sales (RD/S) as R&D expenses (annual item xrd) divided by sales (item sale). Note that the number of stocks for which this factor is available during some of the first 22 investment periods is less than 500, and the number available from the 23th investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 23$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.6 RD/M

As in Chan et al. (2001), we measure R&D-to-market (RD/M) as R&D expenses (annual item xrd) for calendar year $t - 1$ divided by the market capitalization (ME) on day $h - 1$, where t denotes the current calendar year. Note that the number of stocks for which this factor is available during some of the first 22 investment periods is less than 500, and the number available from the 23th investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 23$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.7 Rc/A

Following Li (2011), we measure R&D capital-to-assets (Rc/A) as the ratio of R&D capital (Rc) to total assets (item at). Rc is a weighted average of R&D expenses (annual item xrd) over the last five calendar years with a depreciation rate of 20%:

$$\mathrm{Rc} := \mathrm{xrd}_{t-1} + 0.8\mathrm{xrd}_{t-2} + 0.6\mathrm{xrd}_{t-2} + 0.4\mathrm{xrd}_{t-4} + 0.2\mathrm{xrd}_{t-5} \ , \tag{C.13}$$

where t denotes the current calendar year. Note that the number of stocks for which this factor is available during some of the first 30 investment periods is less than 500, and the number available from the 31st investment date to the end is always less than 1000. As a result, for dimension $N = 500$, we start the portfolio formation on investment date $h = 31$ whereas for dimension $N = 1000$, we do not consider this factor.

C.5.8 OL

Following Novy-Marx (2011), we measure operating leverage (OL) as cost of goods sold (quarterly item cogsq) plus selling, general, and administrative expenses (item xsgaq), then divided by total assets (item atq). Note that the number of stocks for which this factor is available during the first 32 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 33$.

C.6 Trading Frictions

C.6.1 Turn

Following Datar et al. (1998), we measure the share turnover (Turn) as its average daily share turnover over the previous six months from $t-6$ to $t-1$ (that is, from day $h-126$ through day $h-1$). Daily turnover is the number of shares traded (item vol) divided by the number of shares outstanding (item csho). To account for the institutional features of the way NASDAQ and NYSE volume are reported, we adjust the trading volume for NASDAQ stocks as in Gao and Ritter (2010): Previous to 02/01/2001, we divide NASDAQ volume by 2.0; from 02/01/2001 through 12/31/2001, we divide NASDAQ volume by 1.8; for 2002 and 2003, we divide NASDAQ volume by 1.6; and from 2004 on, we use the original NASDAQ volume. Reversal of Turn is used as the actual factor.

C.6.2 Tvol

Following Ang et al. (2006), we measure total volatility (Tvol) as the standard deviation of a stock's daily returns over the previous month $t-1$ (that is, from day $h-21$ through day $h-1$). Reversal of Tvol is used as the actual factor.

C.6.3 Avol

Following Bandyopadhyay et al. (2010), we measure accrual volatility (Avol) as the standard deviation of the ratio of total accruals (TA) to total sales (item saleq) over the previous 16 quarters from quarter $t-16$ to quarter $t-1$, where t denotes the current quarter. TA is defined in their equation (7); the only difference is that we use quarterly data here. Reversal of Avol is used as the actual factor. Note that the number of stocks for which this factor is available during the first 27 investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 28$.

C.6.4 Cvol

Following Huang (2009), we measure cash flow volatility (Cvol) as the standard deviation of cash flow (CF) over the previous 16 quarters from quarter $t-16$ to quarter $t-1$, where t denotes the current quarter. CF is defined as the sum of income before extraordinary items (item ibq), depreciation and amortization expense (item dpq, zero if missing), and the increase in net non-cash working capital (ΔWC in Section C.3.10 with quarterly data). Reversal of Cvol is used as the actual factor. Note that the number of stocks for which this factor is available during the first six investment periods is less than 1000. As a result, for dimension $N = 1000$, we start the portfolio formation on investment date $h = 7$.

Additional References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299.

Balakrishnan, K., Bartov, E., and Faurel, L. (2010). Post loss/profit announcement drift. *Journal of Accounting and Economics*, 50(1):20–41.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2):427–446.

Bandyopadhyay, S. P., Huang, A. G., and Wirjanto, T. S. (2010). The accrual volatility anomaly. Working paper, School of Accounting and Finance, University of Waterloo.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of Financial Economics*, 12(1):129–156.

Bazdrech, S., Belo, F., and Lin, X. (2008). Labor hiring, investment and stock return predictability in the cross section. Working paper, Department of Finance, London School of Economics and Political Science.

Belo, F. and Lin, X. (2012). The inventory growth spread. *Review of Financial Studies*, 25(1):278–313.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *Journal of Finance*, 43(2):507–528.

Boudoukh, J., Michaely, R., Richardson, M., and Roberts, M. R. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance*, 62(2):877–915.

Bradshaw, M. T., Richardson, S. A., and Sloan, R. G. (2006). The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics*, 42(1):53–85.

Brown, D. P. and Rowe, B. (2007). The productivity premium in equity returns. *Available at SSRN 993467*.

Chan, L. K., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, 51(5):1681–1713.

Chan, L. K., Lakonishok, J., and Sougiannis, T. (2001). The stock market valuation of research and development expenditures. *Journal of Finance*, 56(6):2431–2456.

Chandrashekar, S. and Rao, R. K. (2009). The productivity of corporate cash holdings and the cross-section of expected stock returns. *McCombs Research Paper Series No. FIN-03-09*.

Chemmanur, T. J. and Yan, A. (2010). Advertising, attention, and stock returns. Technical report, Boston College. Available at http://ssrn.com/abstract=1340605.

Chincarini, L. B. and Kim, D. (2006). *Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management*. McGraw-Hill, New York.

Chordia, T., Subrahmanyam, A., and Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics*, 59(1):3–32.

Cooper, M. J., Gulen, H., and Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *Journal of Finance*, 63(4):1609–1651.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *Journal of Finance*, 61(4):1605–1643.

Datar, V. T., Naik, N. Y., and Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, 1(2):203–219.

Davis, J. L., Fama, E. F., and French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *Journal of Finance*, 55(1):389–406.

Eberhart, A. C., Maxwell, W. F., and Siddique, A. R. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *Journal of Finance*, 59(2):623–650.

Fama, E. F. and French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51(1):55–84.

Foster, G., Olsen, C., and Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *Accounting Review*, 59(4):574–603.

Gao, X. and Ritter, J. R. (2010). The marketing of seasoned equity offerings. *Journal of Financial Economics*, 97(1):33–52.

Gettleman, E. and Marks, J. M. (2006). Acceleration strategies. *SSRN Working Paper Series*.

Green, J., Hand, J. R., and Zhang, F. (2014). The remarkable multidimensionality in the cross-section of expected us stock returns. *Available at SSRN 2262374*.

Gu, L., Wang, Z., and Ye, J. (2009). Information in order backlog: change versus level. Working paper, Citeseer.

Hafzalla, N., Lundholm, R., and Matthew Van Winkle, E. (2011). Percent accruals. *Accounting Review*, 86(1):209–236.

Haugen, R. A. and Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3):401–439.

Hirshleifer, D., Hou, K., Teoh, S. H., and Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, 38:297–331.

Houge, T. and Loughran, T. (2000). Cash flow is king? Cognitive errors by investors. *Journal of Psychology and Financial Markets*, 1(3-4):161–175.

Huang, A. G. (2009). The cross section of cashflow volatility and expected stock returns. *Journal of Empirical Finance*, 16(3):409–429.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance*, 49(5):1541–1578.

Lerman, A., Livnat, J., and Mendenhall, R. R. (2008). The high-volume return premium and post-earnings announcement drift. *Available at SSRN 1122463*.

Li, D. (2011). Financial constraints, R&D investment, and stock returns. *Review of Financial Studies*, 24(9):2974–3007.

Litzenberger, R. H. and Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics*, 7(2):163–195.

Lyandres, E., Sun, L., and Zhang, L. (2008). The new issues puzzle: Testing the investment-based explanation. *Review of Financial Studies*, 21(6):2825–2855.

Novy-Marx, R. (2010). The other side of value: Good growth and the gross profitability premium. Technical report, National Bureau of Economic Research.

Novy-Marx, R. (2011). Operating leverage. *Review of Finance*, 15(1):103–134.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, pages 109–131.

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, pages 1–41.

Pontiff, J. and Woodgate, A. (2008). Share issuance and cross-sectional returns. *Journal of Finance*, 63(2):921–945.

Prakash, R. and Sinha, N. (2013). Deferred revenues and the matching of revenues and expenses. *Contemporary Accounting Research*, 30(2):517–548.

Richardson, S. A., Sloan, R. G., Soliman, M. T., and Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3):437–485.

Rosenberg, B., Reid, K., and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, 11(3):9–16.

Scott, A., Sloan, R., Soliman, M., and Tuna, I. (2005). Accrual reliability, earnings persistence and stock returns. *Journal of Accounting Research*, 39:437–485.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? (Digest summary). *Accounting Review*, 71(3):289–315.

Soliman, M. T. (2008). The use of dupont analysis by market participants. *Accounting Review*, 83(3):823–853.

Thomas, J. and Zhang, F. X. (2011). Tax expense momentum. *Journal of Accounting Research*, 49(3):791–821.

Thomas, J. K. and Zhang, H. (2002). Inventory changes and future returns. *Review of Accounting Studies*, 7(2-3):163–187.

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39(04):677–700.

Xing, Y. (2008). Interpreting the value effect through the q-theory: An empirical investigation. *Review of Financial Studies*, 21(4):1767–1795.