



Characteristics-based portfolio choice with leverage constraints[☆]



Manuel Ammann^{a,1}, Guillaume Coqueret^{b,2}, Jan-Philip Schade^{a,*}

^a University of St.Gallen, Rosenbergstrasse 52, 9000 St.Gallen, Switzerland

^b Montpellier Business School, 2300 Avenue des Moulins, 34080 Montpellier, France

ARTICLE INFO

Article history:

Received 4 March 2015

Accepted 21 April 2016

Available online 11 June 2016

JEL classification:

G11

Keywords:

Portfolio choice

Leverage constraint

Characteristics-based investing

ABSTRACT

We show that the introduction of a leverage constraint improves the practical implementation of characteristics-based portfolios. The addition of the constraint leads to significantly lower transaction costs, to a reduction of negative portfolio weights, and to a decrease in volatility and misspecification risk. Furthermore, it allows investors to implement any desired level of leverage. In this study, we include 12 characteristics, thereby extending the classical size, book-to-market and momentum paradigm. We report several key indicators such as the proportion of negative weights, Sharpe ratio, volatility, transaction costs, the transaction cost-adjusted certainty equivalent returns, and the Herfindahl–Hirschman index. Analyzing the sensitivity of these key indicators to the choice of multiple combinations of the 12 characteristics, to risk aversion, and to estimation sample size, we show that constrained policies are much less sensitive to these parameters than their unconstrained counterparts. Finally, for quadratic utility, we derive a semi-closed analytical form for the portfolio weights. Overall, we provide a comprehensive extension of characteristics-based portfolio choice and contribute to a better understanding and implementation of the allocation process.

© 2015 Published by Elsevier B.V.

1. Introduction

Fama and French (1992) show that the size and book-to-market ratio of companies are strong drivers of the cross-sectional differences in stock returns, and they also document that the performance of stocks can be explained by exposures to factors based on these drivers (in addition to the aggregate market). The most intuitive way to create a portfolio with exposure is to select stocks with the desired attributes. Recently, new methods have been proposed. These techniques allocate wealth according to synthetic measures, as in Walkshäusl and Lobe (2010), Arnott et al. (2005) and Asness et al. (2013), or rely on more systematic methods based on optimization procedures, as in Brandt et al. (2009), Hand and Green (2011), Hjalmarsson and Manchev (2012) and

Boudt et al. (2014). However, those optimization schemes may not be practical because they imply high levels of leverage and large transaction costs. We therefore propose an extension to those techniques that facilitates its implementation and enables investors to dynamically adapt their leverage over time. Additionally, we provide new insights into the mechanics of the model with respect to risk aversion, sample size, and constraint intensity.

The idea of combining firm characteristics with systematic optimization procedures brings several advantages for investors. One of the main practical benefits of focusing on characteristics in optimization schemes is the reduction in dimensionality when the investment universe is large. For instance, a classical minimum variance optimization for the S&P 500 universe requires the computation of more than 125,000 covariances. Furthermore, when basing their investment decisions on risk factors, investors can transfer their beliefs in different return drivers directly into their portfolio weights. Consequently, such approaches enable each investor to create investment strategies according to his or her taste. For this purpose, an intuitive framework was proposed by Brandt et al. (2009): building on the insights of Daniel and Titman (1997, 1998), they suggest modeling optimal portfolio weights by deviating from an initial benchmark using a linear function of normalized characteristics. This is achieved by optimizing the expected

[☆] The authors would like to thank Alexandre Baptista, Jens Jackwerth, Abraham Lioui, two anonymous referees and the participants of the Finance Seminar at the University of Constance, the European Financial Management Annual Meeting and the Conference of the French Finance Association for their helpful comments. All mistakes remain ours.

* Corresponding author. Tel.: +41 (0)71 224 70 80.

E-mail addresses: manuel.ammann@unisg.ch (M. Ammann), g.coqueret@montpellier-bs.com (G. Coqueret), jan-philip.schade@unisg.ch (J.-P. Schade).

¹ Tel.: +41 (0)71 224 70 80.

² Tel.: +33 (0)467 10 26 78.

utility of the future wealth with respect to the loadings of the characteristics.

Nevertheless, this method entails limitations when applying it in practice. Similar to the classical [Markowitz \(1952\)](#) mean–variance portfolios, characteristics-based portfolios are usually very leveraged: the optimal solutions imply large negative weights and many stocks must be shorted. For example, Tables 1 and 6 of [Brandt et al. \(2009\)](#) report average sums of negative weights above 100% in their empirical results. In practice, such levels of leverage are often infeasible and they indirectly lead to high asset turnover and large transaction costs: [Brandt et al. \(2009\)](#) display annual turnovers above 100%.³ At the end of their article [Brandt et al. \(2009\)](#) advocate an extension of their model in which they set negative positions to zero and rescale the weights. However, this approach only enables investors to set their leverage completely to zero and, as we find in our study, is only optimal for low levels of risk aversion. The investor might be interested in *limited* levels of leverage, which is, to the best of our knowledge, absent from the literature on characteristics-based investing.⁴

This article has two major objectives. The first is to introduce an easy-to-implement leverage constraint to characteristics-based portfolio optimization to overcome the aforementioned limitations. Therefore, we restrict the original deviation from the initial benchmark to a given L^2 -norm. Such an approach has several virtues. First, it entails reduced levels of leverage. This is notably attractive for investors who are sensitive to margin requirements but also for those who have strong compliance-driven short-selling restrictions.⁵ Second, the constraint lessens the sensitivity of the performance to the selection of firm characteristics and risk aversion. Consequently, the risk of model misspecification is significantly attenuated. Finally, constrained policies are more stable over time compared to unconstrained portfolios, reducing asset rotation and curtailing corresponding transaction costs. Therefore, when adjusting the portfolios' certainty equivalent return by transaction costs, we find that constrained portfolios are economically superior to their unconstrained counterparts. Given that our findings persist in several robustness tests (which include the CRRA utility function), these improvements highlight the benefits investors can expect from the methodology we propose.

Our second objective is to improve our understanding of the performance and sensitivity of the final portfolios with respect to both inputs: the various firm characteristics and the model parameters. The large amount of available firm characteristics contrasts the handful of attributes that have been considered so far. [Brandt et al. \(2009\)](#) argue that since market capitalization, book-to-market ratios and past returns suffice to explain the cross-section of returns (as shown by [Fama and French \(1992\)](#) and [Carhart \(1997\)](#)), they are good enough candidates to be entered in the optimization program. [Hjalmarsson and Manchev \(2012\)](#) use the exact same attributes. While we do not question the relevance of this choice, we adopt a more agnostic approach, which allows for a broader set of

characteristics, both in terms of accounting figures and moments of past returns. We consider a set of 12 characteristics and study their impact on the performance of the portfolio policies. As such, we do not restrict our study to a few combinations of characteristics, but we span, in total, no less than 298 assortments. This analysis can be used by potential investors as a guideline to find and understand which firm characteristics suit them best. Finally, we study the sensitivity of our results to variations in important input parameters such as risk aversion, estimation sample size, and intensity of the leverage constraint. These sensitivity analyses convey additional insights on the mechanics of the model.

We contribute to the existing literature on constraints within the mean–variance framework, as in [Jagannathan and Ma \(2003\)](#), [DeMiguel et al. \(2009a\)](#), [Fan et al. \(2012\)](#), [Jacobs and Levy \(2013\)](#), [Jacobs and Levy \(2014\)](#) and [Coqueret \(2015\)](#) by adding a leverage constraint to characteristics-based investing. So far, to the best of our knowledge, this extension has not been applied to characteristics-based policies before. Our model is close to that of [Brandt et al. \(2009\)](#), even though, as in [Hjalmarsson and Manchev \(2012\)](#), we consider quadratic utility functions in our base-case scenario. As a supplementary benefit, the resulting closed-form solution allows for a direct interpretation of the terms in the formula. Finally, in addition to [Brandt et al. \(2009\)](#), [Hjalmarsson and Manchev \(2012\)](#), [Hand and Green \(2011\)](#) and [Boudt et al. \(2014\)](#), we broaden the existing empirical evidence on the performance of different firm characteristics within this framework.

The remainder of the paper is structured as follows. In Section 2, we detail our methodology and further justify the inclusion of a constraint in the optimization scheme. In Section 3, we describe our dataset and provide results for portfolios solely based on one or two firm characteristics. Section 4 extends our model to several applications and underlines the flexibility of our constraint. Section 5 is devoted to robustness checks related to the optimal number of characteristics that should be considered, to the sensitivity of our results to sample size and the bindingness of the constraint. Finally, in Section 6, we conclude.

2. Methodology

2.1. The model

Our starting point is the framework introduced by [Brandt et al. \(2009\)](#), who consider policies which take the following linear form⁶:

$$\mathbf{w}_T = \bar{\mathbf{w}}_T + \mathbf{x}_T \boldsymbol{\theta}_T, \quad (1)$$

where $\bar{\mathbf{w}}_T$ is an initial benchmark which is adjusted according to the cross-sectional differences in characteristics. The $(F_T \times 1)$ vector $\boldsymbol{\theta}_T$ is the weight assigned to the characteristics and the $(N_T \times F_T)$ matrix \mathbf{x}_T comprises the firm's characteristics normalized so that they have zero mean and unit variance. N_T henceforth denotes the number of stocks and F_T the number of characteristics at time T . We use bold notations for vectors and matrices. Moreover, we use subscripts to underline that the portfolios are time-dependent: we consider dynamic trading and weights are updated at each rebalancing period. Accordingly, we seek to solve the following max-utility problem:

$$\max_{\boldsymbol{\theta}_T} \mathbb{E}_T[u(r_{p,T+1})] = \max_{\boldsymbol{\theta}_T} \mathbb{E}_T\left[u\left(\left(\bar{\mathbf{w}}_T + \mathbf{x}_T \boldsymbol{\theta}_T\right)' \mathbf{r}_{T+1}\right)\right], \quad (2)$$

³ These figures correspond to a gross leverage above 3. In their study on hedge-fund leverage, [Ang et al. \(2011\)](#) show that gross leverage fluctuated between 1 and 2 for equity hedge funds between 2005 and 2010. Using another database, [Farnsworth \(2014\)](#) finds similar values and shows that leverage remained at even lower levels between 2010 and 2013. Since hedge-funds are among the most levered asset managers, it is unlikely that many investors can realize a leverage of 2 or more.

⁴ [Jacobs and Levy \(2007\)](#), [Lo and Patel \(2008\)](#), [Fan et al. \(2012\)](#) and [Coqueret \(2015\)](#) report that reasonable leverage is likely to increase the risk-adjusted performance of portfolios.

⁵ For instance, in the US, Regulation T of the Federal Reserve Board requires that the sum of the absolute value of all positions does not exceed twice the equity within the account (i.e. the margin requirement is at least equal to 50%). The impact on prices of heterogeneous margin requirements across assets was studied by [Carleanu and Pedersen \(2011\)](#). [Rytchkov \(2014\)](#) considers a similar problem with only one asset which is subject to time-varying margin constraints.

⁶ [Brandt et al. \(2009\)](#) normalize the second term by the number of stocks, but in our framework a scaling simplification occurs in the computation of the optimal $\boldsymbol{\theta}_T$ which cancels this normalization.

where \mathbf{r}_{T+1} is the $(N \times 1)$ vector of the firms future returns and $r_{p,T+1}$ is the aggregate future return of the portfolio. The expectation's underscript T highlights that we take the conditional expectation (the investment decision is taken with knowledge of present and past information only).

In formula (1), we see that the elements of $\mathbf{x}_T \boldsymbol{\theta}_T$ are simply corrections that are applied to the benchmark so as to improve its performance. However, when the magnitude of the corrections is too large, the benchmark weights are diluted and leverage (negative weights) appears. This can be resolved by imposing that none of the weights be negative, similarly as in Jagannathan and Ma (2003). For tractability purposes, we introduce an alternative constraint. We propose performing the maximization program (2) under the constraint

$$\boldsymbol{\theta}_T' \mathbf{x}_T' \mathbf{x}_T \boldsymbol{\theta}_T = \delta_T, \quad (3)$$

where δ_T satisfies the inequalities $(\boldsymbol{\theta}_T^0)' \mathbf{x}_T' \mathbf{x}_T \boldsymbol{\theta}_T^0 > \delta_T > 0$, with $\boldsymbol{\theta}_T^0$ being the solution to the unconstrained problem, so that the constraint is indeed binding. This simply amounts to impose that the L^2 -norm of the vector $\mathbf{x}_T \boldsymbol{\theta}_T$ is equal to δ_T . Of course, one could consider an inequality instead of an equality in the constraint, but given the convex nature of the problem, this would lead to the same solution (the constraint corresponds to the surface of an ellipsoid). As δ_T decreases to zero, the optimal portfolio converges to the initial benchmark. Compared to the unconstrained case, the inclusion of (3) in the optimization reduces the magnitude of the elements of $\mathbf{x}_T \boldsymbol{\theta}_T$ which increases the relative importance of the prior benchmark $\bar{\mathbf{w}}_T$. Accordingly, the intensity of the constraint can be fine-tuned to match any tracking-error target with respect to the benchmark.

In this article, we set the benchmark starting point to be the equally-weighted portfolio: $\bar{\mathbf{w}}_T = \mathbf{1}_{N_T}/N_T$, where N_T is the number of stocks considered by the investor and $\mathbf{1}_N$ is an N -dimensional vector of ones. A popular alternative would be the value-weighted portfolio, but this would set the market capitalization as an important driver of the final weights. We prefer to stick with an agnostic prior, and this choice can be further justified by the fact that the $1/N$ portfolio has been shown to consistently outperform other benchmarks, including the value-weighted portfolios (see DeMiguel et al. (2009b) and Plyakha et al. (2012)). Moreover, Pflug et al. (2012) have proven that when the loss distribution is highly ambiguous, the $1/N$ portfolio becomes optimal. If the benchmark is equally weighted, there is a simple equivalence between the constraint (3) and a constraint on total weights: $\boldsymbol{\theta}_T' \mathbf{x}_T' \mathbf{x}_T \boldsymbol{\theta}_T = \delta_T \Leftrightarrow \mathbf{w}_T' \mathbf{w}_T = \delta_T + N_T^{-1}$ because the characteristics' matrix is normalized and the elements of $\mathbf{x}_T \boldsymbol{\theta}_T$ sum to zero. This is a convenient property because the constraint on the perturbations $\mathbf{x}_T \boldsymbol{\theta}_T$ translates into a constraint on the final weights. Lastly, in the context of leverage constraints, an equally-weighted starting point ensures that all weights are at the same distance from zero, which reduces the odds of negative weights within the cross-section of assets.

2.2. Quadratic utility versus CRRA

Utility functions with constant relative risk aversion (CRRA) are one common approach to model investor behavior in optimization schemes.⁷ One of their key features is that they take higher moments of portfolio returns into account. However, some numerical problems can arise when using CRRA functions in optimization methods. First, as pointed out by Hentschel and Long (2004),

King et al. (2002) and Geweke (2001), some CRRA optimizations can fail to identify the global optimum, whether it be because numerical issues or because the problem is ill-posed. Second, Yoon (2004) shows that for a large class of realistic processes, the expected utility is not even finite under CRRA preferences. Lastly, the choice of CRRA utility in the context of characteristics-based optimization can result in a situation in which it is optimal to take unlimited positive and negative portfolio weights (the solution is unbounded - the proof of this result is available upon request).

These shortcomings are overcome by quadratic utility functions. In contrast to their CRRA peers, they enable closed-form solutions which guarantee a global maxima under very mild assumptions. For this reason, we perform our base case analysis using quadratic utility. However, it is to note that the benefits of the introduced constraint are not dependent on the choice of the chosen utility function. As our robustness checks show, the improvements hold as well when using CRRA utility. The choice of quadratic preferences is not very restrictive: given the low frequency at which accounting figures are released, we use year-on-year returns. One year is a rather long horizon for the computation of returns and, as Campbell et al. (1997) put it, "since all moments are finite, the Central Limit Theorem applies and long-horizon returns tend to be closer to the normal distribution than short-horizon returns". As such, if returns are close to Gaussian, the merits of CRRA functions vanishes.

2.3. The derivation of the constraint tightness and its interpretations

In practice, the vector $\boldsymbol{\theta}_T$ must be estimated using past data. More precisely, we seek the solution of

$$\max_{\boldsymbol{\theta}_T} \frac{1}{T} \sum_{t=T-\tau}^{T-1} u \left(\sum_{i=1}^{N_T} (\bar{\mathbf{w}}_{i,t} + \boldsymbol{\theta}_T' \mathbf{x}_{i,t}) r_{i,t+1} \right), \quad (4)$$

subject to $\boldsymbol{\theta}_T' \mathbf{x}_T' \mathbf{x}_T \boldsymbol{\theta}_T = \delta_T$,

where $t = T$ is the present date and $t = T - \tau$ is the first date of the estimation sample. For a vector \mathbf{x} , we write x_i its i^{th} element and for a matrix \mathbf{X} , \mathbf{X}_i denotes its i^{th} column. We do not need to impose that the final weights sum to one because the linear form (1) and the demeaning of \mathbf{x}_T ensure that it is the case. For implementation purposes, at a given date, the sample has a constant number of stocks (N_T) over all past years: the optimization is performed only on stocks for which the characteristics are available from date $t = T - \tau$ to date $t = T - 1$. In this setting, the sample size is equal to τ and we impose that $\tau > F_T$ and $N_T > F_T$.⁸

Proposition 2.1. Under the assumption of a quadratic utility function $u(x) = x - \gamma x^2/2$, the solution of (4) is equal to $\boldsymbol{\theta}_T^*(\lambda^*)$, where

$$\boldsymbol{\theta}_T^*(\lambda) = [2\lambda^* \boldsymbol{\Sigma}_T^{(x)} + \gamma \boldsymbol{\Sigma}_T^{(p)}]^{-1} \times [\boldsymbol{\mu}_T - \gamma \boldsymbol{\sigma}_T^{(\bar{\mathbf{w}})}], \quad (5)$$

where $\lambda^* = \inf\{\lambda > 0, (\boldsymbol{\theta}_T^*(\lambda))' \mathbf{x}_T' \mathbf{x}_T \boldsymbol{\theta}_T^*(\lambda) = \delta_T\}$.

This representation recalls those based on regressions of Hjalmarsen and Manchev (2012), except that new terms appear because of the constraint and the benchmark portfolio:

$$\boldsymbol{\mu}_T = \frac{1}{T} \sum_{t=T-\tau}^{T-1} \mathbf{x}_t' \mathbf{r}_{t+1}, \quad \boldsymbol{\Sigma}_T^{(x)} = \mathbf{x}_T' \mathbf{x}_T, \quad \boldsymbol{\sigma}_T^{(\bar{\mathbf{w}})} = \frac{1}{T} \sum_{t=T-\tau}^{T-1} \mathbf{x}_t' \mathbf{r}_{t+1} \mathbf{r}_{t+1}' \bar{\mathbf{w}}_t,$$

and $\boldsymbol{\Sigma}_T^{(p)} = \frac{1}{T} \sum_{t=T-\tau}^{T-1} \mathbf{x}_t' \mathbf{r}_{t+1} \mathbf{r}_{t+1}' \mathbf{x}_t,$ (6)

⁷ For example Brandt et al. (2009) use CRRA functions for all the calculations in their empirical analysis.

⁸ As is shown in the proof, these conditions ensure that the solution of the problem exists.

we can interpret the terms in Eq. (6) in the following way: first, the vectors $\mathbf{x}'_t \mathbf{r}_{t+1}$ correspond to the $(F_T \times 1)$ returns of portfolios with weights \mathbf{x}_{iT} ($i = 1, \dots, F_T$) and the $(F_T \times 1)$ vector $\boldsymbol{\mu}_T$ carries their past average values. Likewise, $\boldsymbol{\Sigma}_T^{(P)}$ is equal to the $(F_T \times F_T)$ sample covariance matrix of the characteristics-based portfolios inferred from their past returns. It is nonsingular as long as $\tau > F_T$. The scaled instantaneous covariance matrix of the time- T characteristics ($\boldsymbol{\Sigma}_T^{(x)}$) is invertible when $N_T > F_T$. Lastly, the vector $\boldsymbol{\sigma}_T^{(\bar{\mathbf{w}})}$ measure the covariance between the benchmark portfolio and the characteristics-weighted portfolios. Therefore, in the unconstrained case ($\lambda^* = 0$), the optimal parameter $\boldsymbol{\theta}_T^*(\lambda^*)$ in Eq. (5) can be decomposed in two components: the first is equal to the maximum Sharpe portfolio where the assets are the characteristics-based portfolios, and the second is an adjustment stemming from the covariance with the benchmark starting point.⁹

Next, we discuss the mechanics of the optimization embedded in Eq. (5). First, when the past returns and firm characteristics are given, λ^* is entirely driven by δ_T and as δ_T decreases to zero, the constraint becomes more binding and λ^* increases to infinity. Mechanically, the magnitude of the values of $\boldsymbol{\theta}_T^*$ decline, which is the sought effect. In the limit δ_T decreases to zero, λ^* increases to infinity and of course $\boldsymbol{\theta}_T$ goes to zero as well. More technically, the constraint (3) acts like a regularization à la Tikhonov: if the sample size is too small ($\tau \leq F_T$), the unconstrained problem is ill-conditioned, but adding the constraint guarantees the existence and uniqueness of a solution, as long as the number of assets exceeds the number of characteristics (which is always the case in practice).¹⁰

The second important variable in Eq. (5) is the risk aversion parameter. In the second factor of the product, it is straightforward that the relative importance of the past average values $\boldsymbol{\mu}_T$ decreases when risk aversion increases. In fact, when γ increases to infinity, the solution converges to

$$\boldsymbol{\theta}_T = - \left[\sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \mathbf{x}_t \right]^{-1} \times \left[\sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \bar{\mathbf{w}}_t \right],$$

and this expression is expected to generate a low variance portfolio because infinite risk aversion corresponds to a utility function that focuses only on the quadratic term.

Going back to the simple form (1), we see that the policy is a combination of a benchmark portfolio plus F_T portfolios (one for each characteristic) with weights equal to the elements of $\boldsymbol{\theta}_T$. When (2) is unconstrained, the magnitudes of the elements of $\boldsymbol{\theta}_T$ are such that the benchmark is diluted in the characteristics portfolios. However, with the introduction of the constraint (3), the weights are progressively shrunk towards the benchmark portfolio as the constraint becomes tighter. Thus, imposing a strong constraint is only efficient if the benchmark is well chosen.

2.4. Choosing δ_T

The aim of the constraint (3) is to reduce the impact of the adjustment to the benchmark. As such, the choice of δ_T determines to

⁹ If N assets have expected returns vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, then a standard result of the mean-variance framework is that the maximum Sharpe ratio portfolio is $\{\arg\max_{\mathbf{w}} \frac{\mathbf{w}'\boldsymbol{\mu}}{\sqrt{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}}, \mathbf{1}'\mathbf{w} = 1\} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$.

¹⁰ Formally, the quadratically constrained quadratic program (4) is equivalent, for some real number λ , to the penalized quadratic program

$$\min_{\boldsymbol{\theta}_T} \|\mathbf{A}\boldsymbol{\theta}_T - \mathbf{b}\|_2^2 + \lambda \|\boldsymbol{\theta}_T\|_{\boldsymbol{\Sigma}_T^{(x)}}^2,$$

where $\mathbf{A} = (\frac{\gamma}{2} \boldsymbol{\Sigma}_T^{(P)})^{1/2}$ and $\mathbf{b} = (\frac{\gamma}{2} \boldsymbol{\Sigma}_T^{(P)})^{-1/2} (\gamma \boldsymbol{\sigma}_T^{(\bar{\mathbf{w}})} - \boldsymbol{\mu}_T)$ and $\|\mathbf{X}\|_F^2 = \mathbf{X}'\mathbf{X}$. This program is a generalized ridge regression and the regularization intensity, λ , is entirely determined by δ_T .

what extent the final weights can differ from those of the equally-weighted starting point. We want to determine a heuristic method to set a threshold which generates weights significantly different from the extreme cases (zero and full constraint) and simultaneously reduce the proportion of negative weights so that the leverage of the portfolio reaches realistic levels. We note y_{iT} for the elements of the vector $\mathbf{x}_T \boldsymbol{\theta}_T$ and hence the constraint (3) reads

$$\sum_{i=1}^{N_T} y_{iT}^2 = \delta_T. \quad (7)$$

For any given time T , the distribution of the y_{iT} is difficult to identify in general because it depends on the characteristics which enter the optimization and also on the signs of the elements of $\boldsymbol{\theta}_T$. For simplicity and tractability, we assume that the y_{iT} are i.i.d. and normally distributed with zero mean and variance equal to δ_T/N_T (given by (7)). This simply means that there are many small deviations from the benchmark and a few large ones. Then, we have the following result.

Lemma 2.1. *If the y_{iT} are i.i.d. and follow a $N(0, \delta_T/N_T)$ distribution, then the average time- T gross leverage of the portfolio is given by*

$$\mathbb{E}[L_T] = \nu(\delta_T N_T) = 1 + \sqrt{\frac{2\delta_T N_T}{\pi}} e^{-(2\delta_T N_T)^{-1}} - 2\Phi(-(\delta_T N_T)^{-1/2}), \quad (8)$$

where Φ is the cumulative distribution function of the Normal law.

The lemma shows that the average gross leverage in this case depends solely on the product $\delta_T N_T$. Moreover, we also show in the proof of the lemma that this average leverage is strictly increasing in $\delta_T N_T$. In Fig. D.4 we plot this leverage as a function of $\delta_T N_T$. We see that the leverage is close to 1.2 for $\delta_T N_T = 1$. In their study on hedge fund leverage, Ang et al. (2011) show that after the 2008 crisis, the average gross leverage for equity funds evolved to between 1.2 and 1.3. With a more recent database, Farnsworth (2014) also reports that leverage remained at moderate levels between 2008 and 2014. Consequently, notwithstanding the simplification of Gaussianity, setting $\delta_T = N_T^{-1}$ would imply average leverages close to those observed in recent periods for the equity hedge fund industry.¹¹ Note that given the range of N_T (>500), any minor change on δ_T (e.g. taking $\delta_T = 1.3/N_T$ instead of $\delta_T = 1/N_T$) will have no impact on the results. Significant changes can only be obtained with truly different values of δ_T (a variation by a factor 5 or 10 at least). This enables us to set a base case value for δ_T with which we will work in the next section. Nevertheless, we will consider other values of δ_T in Section 4 and in Section 5. Overall, we find that our main results do not change when varying δ_T .

3. Empirical analysis

3.1. Data

The construction of the research universe is detailed in Appendix B and comprises 10,303 companies from 1964 to 2013. In total, the number of companies within our sample grows over time from 1,353 (in 1969) to a maximum of 2,652 (in 2003). In order to construct characteristics-based portfolios we have calculated 12 different firm characteristics for each of the sample companies. A list of all characteristics and their abbreviations can be found in Appendix A. The 12 characteristics do not carry redundant information: Panel A of Table 1 provides the average cross-correlations

¹¹ In unreported results, we find that switching to a Student distribution with 3 degrees of freedom gives strikingly close results. Therefore, our conclusion holds even if the y_{iT} are assumed to be heavy-tailed, i.e. if the large deviations from the benchmark are more frequent.

Table 1

Cross-correlation and auto-correlation of firm characteristics. The table displays the correlation of the firm characteristics within the cross-section and across time. A list of all abbreviations can be found in [Appendix A](#). *Char* refers to the corresponding firm characteristic: *MEQ* is the company's market equity value, *BTM* represents the book-to-market value, *DIY* represents the company's current dividend yield, *LEV* the leverage-ratio, *MOM* the momentum calculated from $t - 12$ to $t - 2$ and *VAR* the variance based on 60 monthly simple returns. *ROA* represents the return on assets, *CFA* the company's cash-flow over assets and *GMV* the absolute annual variation in gross-margin, *ERV* the earnings volatility measured as standard deviation over the past 20 quarters previous to each regarded year and *CUE* the annual change in earnings. Finally, *AGR* stands for the company's year-over-year asset growth. The autocorrelation is computed as follows. Each year, we compute the Pearson correlation of one firm characteristic across all firms with the same characteristic lagged by one to five years. The correlation is calculated only for the firms with available data and then averaged over all sample dates.

Char	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR
<i>Panel A: Cross-correlations</i>												
MEQ	1	−0.13	0.07	−0.08	0.08	−0.17	0.12	0.10	0.03	−0.12	0.00	0.06
BTM		1	0.32	0.48	−0.23	−0.11	−0.21	−0.02	−0.11	−0.13	−0.02	−0.20
DIY			1	0.13	−0.11	−0.40	0.01	0.11	−0.09	−0.12	−0.01	−0.14
LEV				1	−0.12	0.04	−0.23	−0.07	−0.08	0.09	−0.02	−0.11
MOM					1	−0.03	0.07	0.11	0.05	−0.04	0.02	0.01
VAR						1	−0.14	−0.15	0.05	0.30	0.00	0.04
ROA							1	0.52	0.17	−0.25	0.04	0.22
CFA								1	0.07	−0.17	−0.01	−0.12
GMV									1	0.00	0.02	0.09
ERV										1	−0.01	0.00
CUE											1	0.06
Lag	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR
<i>Panel B: Autocorrelations</i>												
1	0.88	0.73	0.69	0.77	−0.05	0.89	0.64	0.38	0.61	0.82	0.04	0.26
2	0.80	0.60	0.60	0.63	−0.06	0.78	0.51	0.30	0.31	0.70	0.03	0.14
3	0.74	0.51	0.55	0.54	−0.03	0.66	0.45	0.26	0.07	0.58	0.03	0.09
4	0.68	0.44	0.52	0.46	−0.02	0.56	0.41	0.24	−0.17	0.46	0.02	0.07
5	0.63	0.39	0.49	0.40	−0.03	0.45	0.36	0.21	−0.09	0.39	0.01	0.05

between all reported characteristics.¹² The highest observed correlation is 52%, between the cash-flow over assets (*CFA*) and the return on assets (*ROA*).

In Panel B of [Table 1](#), we have computed the autocorrelation of the characteristics with lags equal to 1, 2, 3, 4 and 5 years. Autocorrelation is computed as the average Pearson correlation of one firm characteristic on the cross-sectional level lagged by 1 to 5 years. We note that the firm characteristics are quite stable in the cross-section, except for momentum (*MOM*) and the gross margin variation (*GMV*). In contrast, the market equity (*MEQ*) and variance (*VAR*) are the most stable attributes (large firms remain large and low risk firms remain low risk, at least in relative terms).

3.2. Portfolio construction and key indicators

At the beginning of each year, we compute the portfolio policy according to (1) and (5). Within the scope of the single and double characteristic portfolios we consider two cases. The first one is defined with $\delta_T = \infty$ implying that no constraint is used for the optimization (*unconstrained policy*). This case is analog to the case of [Brandt et al. \(2009\)](#) in which every level of leverage can be attained by the optimization. The second case is set with $\delta_T = N_T^{-1}$ allowing for an intermediate level between no constraint and maximum constraint (*constrained policy*).¹³ For each of these two cases, we look at the impact of risk aversion on the performance of the portfolio. Accordingly, we report results for low risk aversion ($\gamma = 1$), moderate risk aversion ($\gamma = 5$) and high risk aversion ($\gamma = 10$).

The portfolio is held for one year and the weights are then updated using the latest data at the end of June. This approach assures

that every year only the accounting data that is actually available is used, which assures that our results do not contain any insight bias. Further, since the final portfolio weights are always calculated only up to the date of portfolio construction, the reported results show the out-of-sample performance of our constraint. Further, portfolio returns are calculated on a monthly basis as simple returns. As in [Brandt et al. \(2009\)](#), we use a calibration sample of $\tau = 10$ years. We study the sensitivity of the results to variations in τ in Section 5.2. We truncate our sample before 1969 in order to obtain a sufficient number of companies for which many characteristics are available. Consequently, the allocation starts in June 1980 and ends in June 2013.

In order to exploit the effect of our constraint on performance, leverage and turnover we report six different measures. The first is the average of transaction costs (*TC*). These are modeled according to the same cross-sectional distribution as that of [Brandt et al. \(2009\)](#): $z_{i,T} = A_T(0.006 - 0.0025me_{i,T})$, where $me_{i,T}$ is the time T market equity of stock i , divided by the time T maximum market equity across all stocks. The A_T factor is used to model a linear decrease of transaction costs in time. However, because our sample is longer than that of [Brandt et al. \(2009\)](#), we assume that transaction costs in 1980 are five times larger than those in 2013 (i.e. $A_{1979} = 5$ and $A_{2013} = 1$). Hence, the transaction cost incurred by stock i at time T is $TC_{i,T} = z_{i,T} \times |w_{i,T} - w_{i,T-}|$, where $w_{i,T-}$ is the weight in portfolio of stock i just before the rebalancing. We compute the average annual portfolio transaction costs as

$$TC = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{N_t} TC_{i,t}.$$

Turning to performance indicators, we compute volatility (*Vol*) as the standard deviation of annual returns. The Sharpe ratio (*SR*) is equal to the annualized return minus the risk free rate, divided by the portfolio's annualized volatility.¹⁴ In all tables, we test whether the Sharpe ratio of a portfolio policy is significantly above that of the equally-weighted benchmark using the bootstrap test of [Ledoit and Wolf \(2008\)](#).

¹² Originally, we assessed 16 different firm characteristics but decided to exclude four of them from our study because they were too correlated with one or more of the 12 we propose. In detail, we have also tested the variance of the returns over the past 24 months and idiosyncratic risk (standard deviation of the residuals of the CAPM regression used for the computation of the beta), but they yielded results too close to those of the 60 month variance. Moreover, we found that the ratio of EBIT to *MEQ* had a 76% correlation with the *MOM* characteristic. Consequently, these attributes were withdrawn from our study.

¹³ Note that we test other values of δ_T in our robustness checks.

¹⁴ We use the 3 M T-Bill rates, which, over the whole sample, average to a 5.6% annual rate.

Figures are presented in bold font when the corresponding p -value is smaller than 10%. While the Sharpe ratio is the most natural performance indicator (overwhelmingly used by practitioners and in the academic literature), we also choose to quantify the economic gain of the leverage constraint through the TC-adjusted certainty equivalent excess return (CER).¹⁵ As in Goto and Xu (2015), we define it by

$$\text{CER} = \hat{\mu} - \frac{\gamma}{2} \hat{\sigma}^2 - \text{TC}, \quad (9)$$

where $\hat{\mu}$ is the average annual return of the portfolio, $\hat{\sigma}^2$ its annualized realized volatility and γ the individual investor's risk aversion. The CER can be interpreted as the increase in the risk-free rate for which the investor would be indifferent between holding cash and holding the risky portfolio, after accounting for transaction costs. Moreover, in our framework, it can be viewed as an estimation of realized utility, net of transaction costs.

In order to quantify the leverage of the portfolios, we provide the average proportion of negative weights (PNW):

$$\text{PNW} = \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^{N_t} \frac{\mathbf{1}_{\{w_{i,t} < 0\}}}{N_t}.$$

Lastly, we assess the diversification of the portfolios through the average Herfindahl–Hirschman index (HHI) – here we follow Goetzmann and Kumar (2008):

$$\text{HHI} = \frac{1}{T} \sum_{t=1}^T \mathbf{w}_t' \mathbf{w}_t.$$

The L^2 -norm of the weights is bounded below by $1/N$, which is the score of the equally-weighted portfolio (highest diversification). It increases to infinity as the amplitude of leverage rises. Consequently, a low HHI signals a high degree of portfolio diversification.

Table 2 provides the results of the introduced key indicators which are obtained for the equally-weighted policy. Its CER is equal to 0.10 for $\gamma = 1$, 0.04 for $\gamma = 5$ and -0.03 for $\gamma = 10$.

3.3. Results for single characteristic portfolio policies

Table 3 shows the results of all single characteristic portfolio policies. The first group of columns of Panel A displays the absolute values of the Sharpe ratio for the unconstrained policies, the Sharpe ratios of the constrained case and the ratios of the constrained to the unconstrained Sharpe ratio (SR ratio). Within the ratio case, a figure smaller one indicates a decrease in the indicator when switching from unconstrained to constrained optimization. The last three columns show the difference in CER when imposing our constraint compared to the unconstrained case (ΔCER). Here, a positive figure indicates that the CER has improved when applying the constraint. As for the SR ratio we report the corresponding ratios for volatility, transaction costs, proportion of negative weights and HHI in Panel B. A list of all abbreviations is provided in Appendix A.

Given the purpose of the constraint, we expect the overall levels of leverage and transaction costs to decrease. Nevertheless, we do not anticipate a monotonous impact on Sharpe ratios when adding the constraint. Also, we expect that the loss of optimality implied by the constraint may be mitigated by the reduction of estimation errors, as in Jagannathan and Ma (2003).

First, we analyze the impact of our constraint on the policies' performance measured by Sharpe ratio and ΔCER . When looking at the unconstrained Sharpe ratios in Panel A of Table 3, we find

Table 2

Key indicators of the benchmark portfolio. This table displays the results for all key indicators for the equally-weighted benchmark portfolio. A list of all abbreviations can be found in Appendix A.

	SR	Vol	TC (%)	PNW	HHI
Equally-weighted benchmark	0.387	0.173	0.549	0.000	0.002

that both the choice of risk aversion and the choice of the underlying characteristic have a significant impact on the resulting out-of-sample performance. For example, the Sharpe ratio decreases by more than 90% for MEQ when comparing $\gamma = 1$ to 10. The cross-sectional differences for all characteristics for $\gamma = 5$ lie between 0.30 and 0.50. In contrast, when introducing our constraint, this divergence is remarkably reduced: sticking to the example of MEQ, the Sharpe ratio decreases by 70% instead of 90% when increasing risk aversion. Further, when regarding the cross-section of all characteristics, the Sharpe ratios for $\gamma = 5$ now lie between 0.34 and 0.45 which, again, shows that the performance relies much less on inputs of the optimization when using the constraint. This contraction in dispersion has a straightforward explanation: the constraint imposes that final weights do not fluctuate too far from the benchmark weights. As a by-product, this also implies a greater stability of weights through time, as shown by the reduction of turnover and transaction costs. We do not observe an unconditional impact of the constraint on the overall levels of Sharpe ratio. This finding also holds when comparing it to the results of the ΔCER (last three rows of Panel A). Surprisingly, the number of policies significantly outperforming the benchmark is higher after the introduction of the constraint (6) than before (4 policies stand out).¹⁶ This is a by-product of the reduction of risk for all policies.

With respect to Panel B of Table 3, we find that the constraint significantly impacts the volatility, transaction costs, proportion of negative weights and HHI. The impact on volatility and HHI is stronger for lower risk aversions than for high ones. For example, the decrease in volatility is 65% for AGR for $\gamma = 1$ whereas it is only 6% when $\gamma = 10$. However, the gain in transaction costs and proportion of negative weights is noteworthy: in 50% of all cases the transaction costs are cut by half and we find that the proportion of negative weights is even curtailed by 70% on average. These findings show that constrained policies are not only easier but also better diversified and cheaper to implement for asset managers with leverage restrictions.

Given the strong impact of the constraint on volatility, transaction costs and leverage and the rather mixed effect on Sharpe ratios, it appears appealing to take a closer look at the CER. In Fig. 1 we plot the empirical cumulative density function (cdf) of all combinations for the single (first row), double (second row) and triple (third row) characteristic portfolio policy for all three levels of risk aversion. The gray lines represent the results of the unconstrained cases and the black lines those of the constrained cases. The dash-dot vertical line shows the CER of the equally-weighted portfolio. This illustration allows us to examine the final net effect of the constraint on all possible combinations of firm-characteristics.¹⁷ In the first row of graphs, we find that the constraint is able to cut off a large amount of underperforming unconstrained portfolios. For $\gamma = 1$ the smallest CER for the unconstrained portfolio is -0.23 whereas the smallest one for the constrained cases is 0.25: an investor would therefore always prefer

¹⁵ Note that we do not use a Sharpe ratio discounted by transaction costs in our base case. However, we perform a related robustness check in Section 5.3.

¹⁶ Notwithstanding the large volatilities when $\gamma = 1$. Given that a Sharpe ratio of 0.71 is not found to be significantly different from 0.39, we may conclude that the test is rather conservative.

¹⁷ The first row has 12 points per figure whereas the second and third have 66 and 220, respectively.

Table 3

Comparison of constrained and unconstrained single characteristic portfolio policies. This table displays the results for all key indicators and different risk aversions of the constrained and unconstrained single characteristic based portfolio policy using constraint (3). We report the absolute values for Sharpe ratios as well as the ratios of constrained to unconstrained key indicators (raw values are available upon request). A list of all abbreviations can be found in [Appendix A](#). The calculation is done for all 12 firm characteristics: *MEQ* is the company's market equity value, *BTM* represents the book-to-market value, *DIY* represents the company's current dividend yield, *LEV* the leverage-ratio, *MOM* the momentum calculated from $t - 12$ to $t - 2$ and *VAR* the variance based on 60 monthly simple returns. *ROA* represents the return on assets, *CFA* the company's cash-flow over assets and *GMV* the absolute annual variation in gross-margin, *ERV* the earnings volatility measured as standard deviation over the past 20 quarters previous to each regarded year and *CUE* the annual change in earnings. Finally, *AGR* stands for the company's year-over-year asset growth. We compute the results of the portfolio policy using different risk-aversion values (γ) of 1, 5 and 10. All results are based on $\delta_T = N_T^{-1}$. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark.

γ	SR (unconstrained)			SR (constrained)			SR ratio			Δ CER		
	1	5	10	1	5	10	1	5	10	1	5	10
<i>Panel A: Sharpe ratios and CER</i>												
MEQ	1.15	0.42	0.10	0.57	0.40	0.18	0.50	0.95	1.80	−0.31	0.00	0.02
BTM	0.88	0.46	0.24	0.54	0.43	0.26	0.61	0.93	1.08	−0.16	0.00	0.01
DIY	0.24	0.50	0.52	0.40	0.42	0.44	1.67	0.84	0.85	0.07	−0.01	−0.01
LEV	0.65	0.42	0.25	0.47	0.40	0.27	0.72	0.95	1.08	−0.07	0.00	0.01
MOM	0.71	0.37	0.15	0.50	0.36	0.23	0.70	0.97	1.53	−0.02	0.02	0.04
VAR	0.35	0.49	0.44	0.40	0.45	0.42	1.14	0.92	0.95	0.04	0.00	0.00
ROA	0.64	0.40	0.23	0.46	0.39	0.29	0.72	0.98	1.26	−0.05	0.01	0.03
CFA	0.35	0.44	0.38	0.40	0.37	0.35	1.14	0.84	0.92	0.23	0.03	0.02
GMV	0.13	0.30	0.29	0.34	0.34	0.34	2.62	1.13	1.17	0.17	0.03	0.05
ERV	0.31	0.42	0.37	0.45	0.45	0.41	1.45	1.07	1.11	0.04	0.01	0.02
CUE	0.32	0.39	0.39	0.37	0.37	0.37	1.16	0.95	0.95	0.07	0.01	0.01
AGR	0.99	0.50	0.23	0.50	0.44	0.27	0.51	0.88	1.17	−0.14	0.00	0.02
	Vol ratio			TC ratio			PNW ratio			HHI ratio		
γ	1	5	10	1	5	10	1	5	10	1	5	10
<i>Panel B: Vol, TC, PNW and HHI</i>												
MEQ	0.37	1.00	0.94	0.15	0.85	0.63	0.45	0.63	0.00	0.02	0.76	0.55
BTM	0.39	0.95	1.00	0.14	0.88	0.90	0.16	0.50	1.00	0.03	0.77	0.94
DIY	0.73	1.00	1.00	0.20	0.64	0.70	0.23	0.33	0.44	0.08	0.49	0.58
LEV	0.34	0.95	1.00	0.14	0.90	0.88	0.00	0.00	1.00	0.02	0.76	0.85
MOM	0.29	0.86	0.94	0.12	0.66	0.64	0.22	0.44	0.64	0.02	0.57	0.65
VAR	0.37	1.00	1.07	0.21	0.63	0.64	0.07	0.33	0.45	0.04	0.44	0.45
ROA	0.31	0.95	0.94	0.12	0.74	0.59	0.12	0.50	0.38	0.02	0.66	0.42
CFA	0.38	1.13	1.13	0.05	0.20	0.23	0.16	0.21	0.32	0.01	0.07	0.08
GMV	0.52	0.94	0.89	0.09	0.33	0.35	0.14	0.24	0.24	0.01	0.16	0.17
ERV	0.46	1.00	1.00	0.20	0.65	0.63	0.00	0.33	0.50	0.04	0.42	0.34
CUE	0.65	1.00	0.94	0.11	0.38	0.50	0.08	0.25	0.33	0.01	0.10	0.16
AGR	0.35	0.90	0.94	0.08	0.61	0.69	0.21	0.50	0.50	0.01	0.51	0.62

the risky *constrained* portfolio to the safe cash holding alternative. Moreover, [Fig. 1](#) shows that a larger proportion of constrained portfolios are able to beat the benchmark, compared to unconstrained ones. Overall, an investor with no prior preference towards a particular characteristic benefits from adding the leverage constraint. Investors confident in a performing characteristic should prefer the unconstrained optimization only if they have a low risk aversion or if they are absolutely certain that this characteristic will drive excess returns out-of-sample. In sum, we find that the constraint decreases the risk of model misspecification and, additionally, reduces estimation risk.

3.4. Results for double characteristic portfolio policies

Having analysed the performance of the single characteristic policies, we now turn to combinations of two characteristics. Twelve characteristics imply 66 pairs. As it is more convenient to display the results of the most important combinations, we only report the top six and bottom six policies. The ranking is performed according to the Sharpe ratio of constrained policies for the intermediate risk aversion ($\gamma = 5$). The results are gathered in [Table 4](#).

Similarly as for the single characteristic portfolios, Panel A displays the absolute values of the Sharpe ratio, the SR ratio as well as Δ CER. Panel B shows the corresponding ratios for volatility, transaction costs and the proportion of negative weights. With respect to the first three columns of Panel A, we find that the best pairs outrank single characteristic policies in terms of Sharpe ratio (1.45 versus 1.15 for $\gamma = 1$ and 0.73 versus 0.5 for $\gamma = 5$).

This can at least partially be explained by the fact that the set of pairs is larger than the set of unique attributes (66 versus 12). Comparing the top six entries with the bottom six ones, we understand that the higher risk-adjusted performance comes from both, a lower volatility and higher returns. When risk aversion is low or medium, nearly all top Sharpe ratio portfolios are significantly higher than that of the equally-weighted benchmark. Similarly to the single characteristic policies, the introduction of the constraint reduces the disparities, both across levels of risk aversion and across choices of indicators.

Analogously to [Table 3](#), we find in Panel B of [Table 4](#) that the constraint has reduced volatility, transaction costs, the proportion of negative weights and improves diversification. Looking at the second row of graphs in [Fig. 1](#), we see that for $\gamma = 5$ and $\gamma = 10$, the cdf for the constrained policies is almost always below that of the unconstrained policies: there are only a handful of combinations of characteristics that allow the unconstrained program to outperform the constrained policies. For $\gamma = 1$, there is a trade-off: the constraint limits both the upside and the downside potential of the parametric portfolios. It is to note, however, that in 85% of the cases, the constrained policies outperform the benchmark (the proportion goes down to 65% for unconstrained policies).

Overall, we find that the performance of unconstrained characteristics-based policies depends strongly on both the risk aversion parameter and the underlying firm characteristics. Therefore, an investor should be very careful when choosing these crucial inputs. The introduction of the leverage constraint reduces the discrepancies across both dimensions (risk aversion and choice

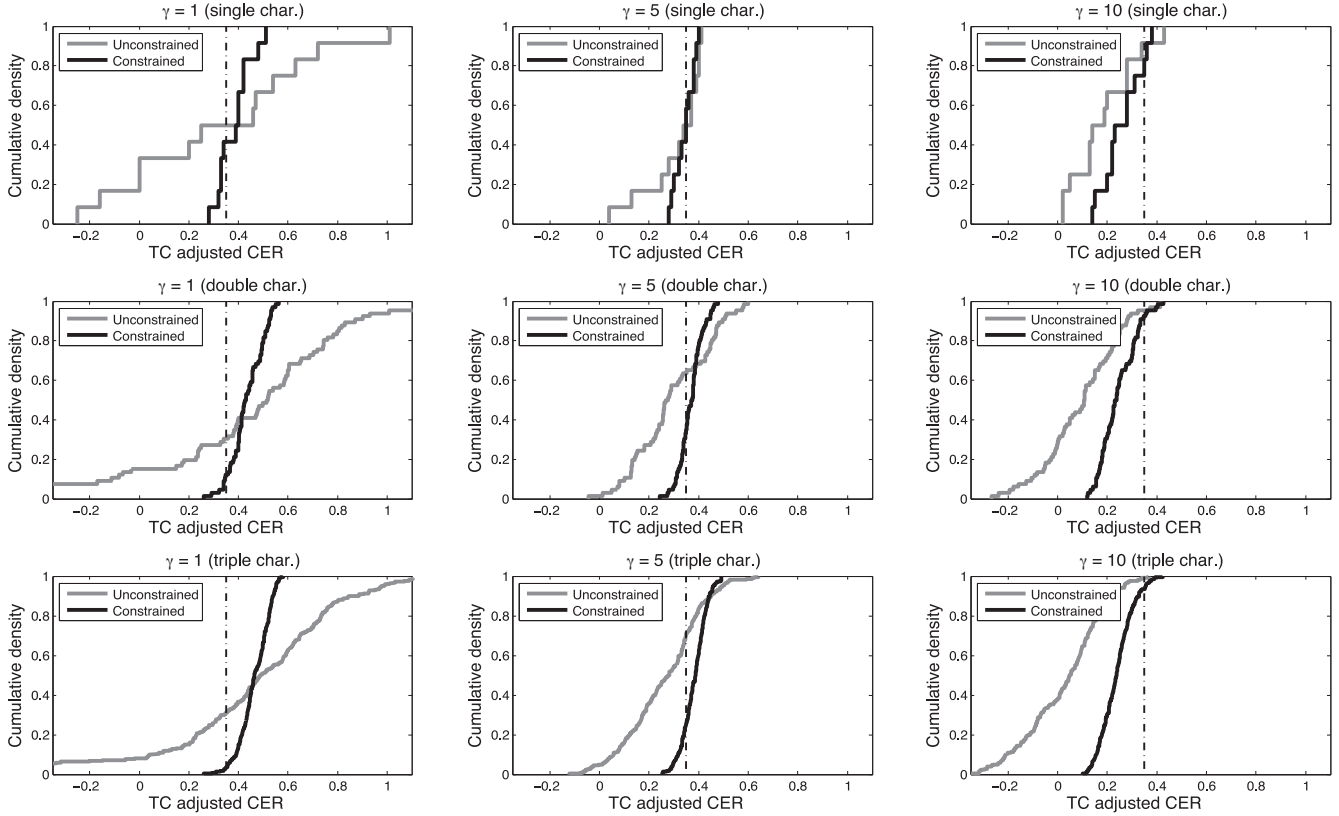


Fig. 1. Empirical distribution of the transaction cost-adjusted CER. This figure shows the cumulative distribution function (cdf) of transaction cost-adjusted certainty equivalent excess return across all combinations of single (first row), double (second row) and triple (third row) characteristic based portfolio policies. The unconstrained policies are plotted in black and the constrained ones in gray. The first column corresponds to $\gamma = 1$, the second to $\gamma = 5$ and the third to $\gamma = 10$. The calibration sample size is $\tau = 10$ and the constraint set at $\delta_T = N_T^{-1}$. The dash-dot line represents the CER of the equally-weighted benchmark.

of characteristic) and hence curtails the risk of ill-advised decisions. Finally, the constraint is again able to improve diversification and reduce the leverage in the double characteristic framework and it diminishes the reported transaction costs.

4. Model extensions

In this section we aim to exploit the high flexibility of our new constraint which enables new applications and model extensions. Given that market equity (MEQ) and the book-to-market ratio (BTM) are the two most frequently cited attributes in the literature, we have decided to rely on a portfolio policy of these two for all empirical analyses in this section.

4.1. Constraining on transaction costs

Instead of constraining the portfolio leverage, the investor might be interested in enforcing an upper bound on the trading costs pertaining to his or her portfolio choice.¹⁸ While recent articles have integrated transaction costs in optimization schemes (e.g. Brown and Smith (2011) and Garleanu and Pedersen (2013)), none of them, to the best of our knowledge, imposes explicit constraints on trading costs.

We recall that time- T transaction costs are computed as $TC_T = \mathbf{z}_T' \Delta \mathbf{w}_T$, where \mathbf{z}_T is the vector of all individual trading costs and $\Delta \mathbf{w}_T = \mathbf{w}_T - \mathbf{w}_{T-}$ gathers all changes in portfolio composition. In order to derive a closed-form solution, we again resort to an L^2 constraint and, similarly to our base case result, we can show that

$$\lambda_T^* = \inf \{ \lambda > 0, \mathbf{z}_T' \times |\bar{\mathbf{w}}_T + \mathbf{x}_T \theta_T^*(\lambda) - \mathbf{w}_{T-}| < tc^* \}$$

applied to (5) leads to the desired upper bound (tc^*) on transaction costs, as long as this bound is reachable. In the equation above, \mathbf{z}_T is the vector of transaction costs detailed in Section 3.2. We provide the values of the key indicators for increasing values of tc^* in Table 5.

We highlight that since the structure of transaction costs evolves through time, tc^* is in fact an average target, computed with prior knowledge of the transaction cost behavior between 1970 and 2012. This explains why there are some discrepancies between tc^* and the actual transaction costs reported in the table. Moreover, when strong changes occur in the investment universe, we allow the constraint to be mildly relaxed so that it can be satisfied.

The results are in line with those of the preceding subsection if we acknowledge that leverage constraints and TC constraints share an almost bijective relationship. A looser TC constraint leads to higher Sharpe ratios but fewer diversified portfolios. Overall, investor welfare (when measured by the CER) decreases when the TC constraint is tight. We highlight that this conclusion could change if the TC penalization embedded in the CER was stronger.

4.2. Time-varying leverage and risk aversion

As shown empirically by Ang et al. (2011) and Farnsworth (2014), the leverage of asset managers is not constant through time. Moreover, leverage is likely to depend on risk-aversion. Indeed, leverage measures the amplitude of the bets within a long/short portfolio and this amplitude is usually smaller in times of uncertainty. Given the flexibility of our constraint to adjust to

¹⁸ We thank an anonymous referee for suggesting this extension of the model.

Table 4

Comparison of constrained and unconstrained double characteristic portfolio policies. This table displays the results for the key indicators and different risk aversions of the constrained and unconstrained single characteristic based portfolio policy using constraint (3). We report the absolute values for Sharpe ratios as well as the ratios of constrained to unconstrained key indicators (raw values are available upon request). A list of all abbreviations can be found in [Appendix A](#). The calculation is done for all 66 pairs. We compute the results of the portfolio policy using different risk-aversion values (γ) of 1, 5 and 10. All results are based on $\delta_T = N_T^{-1}$. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark.

γ		SR (unconstrained)			SR (constrained)			SR ratio			Δ CER		
		1	5	10	1	5	10	1	5	10	1	5	10
Panel A: Sharpe ratios and CER													
top SR	BTM-ERV	0.94	0.73	0.45	0.63	0.54	0.36	0.67	0.75	0.80	−0.20	−0.02	−0.02
	DIY-ERV	0.98	0.71	0.40	0.55	0.53	0.47	0.56	0.75	1.18	0.15	0.02	0.04
	MEQ-VAR	1.36	0.70	0.31	0.60	0.53	0.31	0.44	0.75	1.00	−0.39	−0.02	0.00
	DIY-AGR	1.00	0.69	0.42	0.51	0.52	0.42	0.51	0.75	1.00	−0.15	−0.02	0.02
	MEQ-DIY	1.45	0.69	0.31	0.58	0.51	0.31	0.40	0.74	1.00	−0.41	−0.02	0.00
	VAR-ERV	0.50	0.65	0.57	0.44	0.51	0.49	0.88	0.79	0.86	−0.02	−0.01	0.01
bottom SR	ERV-CUE	0.16	0.32	0.29	0.53	0.37	0.22	3.31	1.17	0.76	0.16	0.02	0.04
	LEV-MOM	0.63	0.32	0.10	0.46	0.37	0.26	0.73	1.16	2.60	0.00	0.05	0.07
	MOM-GMV	0.54	0.31	0.15	0.43	0.36	0.31	0.80	1.17	2.07	0.15	0.06	0.08
	MOM-ROA	0.64	0.31	0.11	0.48	0.35	0.25	0.75	1.13	2.27	0.07	0.06	0.07
	CFA-GMV	0.21	0.30	0.25	0.36	0.34	0.34	1.71	1.16	1.36	0.40	0.05	0.05
	GMV-CUE	−0.02	0.23	0.25	0.33	0.32	0.30	−16.5	1.40	1.20	0.30	0.06	0.08
γ		Vol ratio			TC ratio			PNW ratio			HHI ratio		
		1	5	10	1	5	10	1	5	10	1	5	10
Panel B: Vol, TC, PNW and HHI													
top SR	BTM-ERV	0.34	1.04	1.07	0.14	0.53	0.57	0.17	0.37	0.44	0.02	0.32	0.37
	DIY-ERV	0.30	1.03	1.14	0.04	0.30	0.46	0.10	0.32	0.38	0.01	0.35	0.44
	MEQ-VAR	0.35	1.01	1.14	0.12	0.59	0.62	0.40	0.58	0.47	0.01	0.35	0.47
	DIY-AGR	0.29	0.84	0.88	0.06	0.41	0.38	0.15	0.29	0.39	0.01	0.36	0.37
	MEQ-DIY	0.38	0.94	1.00	0.11	0.53	0.59	0.41	0.40	0.37	0.01	0.41	0.45
	VAR-ERV	0.31	0.94	0.93	0.17	0.41	0.43	0.04	0.30	0.46	0.01	0.22	0.26
bottom SR	ERV-CUE	0.45	1.36	1.20	0.20	0.68	0.56	0.21	0.87	0.78	0.02	0.26	0.18
	LEV-MOM	0.29	0.88	0.89	0.09	0.32	0.31	0.02	0.22	0.31	0.02	0.22	0.25
	MOM-GMV	0.26	0.81	0.89	0.05	0.30	0.32	0.13	0.29	0.29	0.01	0.16	0.20
	MOM-ROA	0.29	0.85	0.89	0.10	0.35	0.45	0.22	0.36	0.50	0.01	0.15	0.23
	CFA-GMV	0.33	0.96	0.94	0.03	0.18	0.21	0.12	0.15	0.21	0.01	0.10	0.11
	GMV-CUE	0.50	0.99	0.90	0.07	0.34	0.37	0.17	0.33	0.33	0.01	0.12	0.13

Table 5

Constraining transaction costs. This table displays the results for all key indicators for policies based on the double characteristic portfolio policy *BTM-MEQ* when the target for transaction costs (tc^*) decreases. A list of all abbreviations can be found in [Appendix A](#). Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark. Transaction costs and HHI are expressed in percent.

tc^*	$\gamma = 1$						tc^*	$\gamma = 5$					
	Vol	SR	TC	PNW	CER	HHI		Vol	SR	TC	PNW	CER	HHI
0.8	0.19	0.45	0.95	0.01	0.11	0.2	0.8	0.18	0.39	0.94	0.03	0.03	0.2
1.2	0.21	0.56	1.33	0.09	0.14	0.3	1.0	0.18	0.41	1.11	0.08	0.04	0.3
1.6	0.22	0.65	1.73	0.17	0.16	0.5	1.2	0.19	0.43	1.26	0.11	0.04	0.3
2.0	0.24	0.72	2.08	0.21	0.18	0.7	1.4	0.19	0.47	1.32	0.15	0.04	0.5
2.4	0.26	0.80	2.46	0.26	0.21	1.0	1.6	0.18	0.49	1.48	0.18	0.05	0.6
2.8	0.28	0.86	2.79	0.29	0.23	1.4	1.8	0.19	0.49	1.52	0.20	0.05	0.6
3.2	0.29	0.89	3.04	0.31	0.24	1.8	2.0	0.18	0.49	1.64	0.19	0.05	0.6

different levels of leverage we extend our model to state-dependent leverage and risk aversion over time.

We proceed with three different state variables which will impact leverage: volatility (volatile markets are expected to reduce investor appetite for leverage), term spread (a positive spread signals economic growth, see [Estrella and Mishkin \(1998\)](#)) and credit spread (hard credit conditions impact firm performance negatively, see [Demchuk and Gibson \(2006\)](#)). Accordingly, we allow both γ and δ_T to depend on one of these variables in the following way:

$$\gamma_t = \gamma \max\left(\frac{v_t}{\bar{v}}, 0.2\right), \quad \delta_t = N_t^{-1} \max\left(\frac{v_t^{-1}}{\bar{v}^{-1}}, 0.2\right), \quad (10)$$

where v is the value of the state variable which is in an increasing bijection with risk aversion (i.e. volatility, credit spread and the inverse of the term spread) and \bar{v} its time-series average. We set a floor of 0.2 so that both leverage and risk aversion remain inside

an interval of $(1/5, 5)$ times the target value. We provide a graphical representation of the time-series of all three state variables in [Fig. 2](#).

Moreover, we test the case where the state variable is equal to the average value of the three variables, once they have been normalized to reach unit mean. The base case indicators summarizing the performance of this extension of our model are collected in [Table 6](#).

In three out of four cases, adjusting both the risk aversion and the level of leverage to macro-economic indicators improves not only the Sharpe ratio, but also the CER. However, it should be noted that the resulting portfolios have a slightly higher proportion of negative weights and are consequently less diversified (their HHI is larger than that of the original method). In unreported results, we find that allowing only the risk aversion or the leverage (but not both) to be time-dependent yielded CERs which were far less competitive. This confirms that leverage should be a function of risk aversion and not set alone by itself.

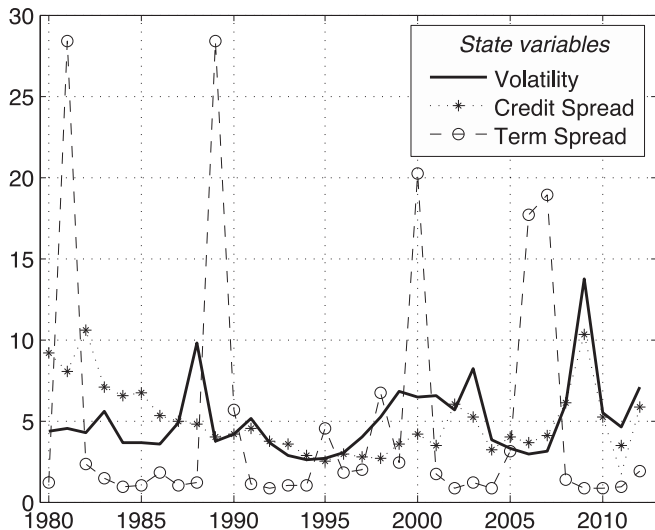


Fig. 2. State variables for time-varying leverage and risk aversion. Volatility is computed as the standard deviation of daily returns over the past year. Term spread is the difference between the 30Y and the 3M US Treasury constant maturity yields. Credit spread is equal to the Moodys Baa minus Aaa corporate bond rates. All variables are normalized so that their average equals one. We only display values from 1980 onwards because this is when the allocation process begins.

4.3. Does adding characteristics increase value?

Scrutinizing all possible combinations of characteristics and picking the best one for a given criterion amounts to pure data snooping. We therefore take a broader view and make use of the numerous combinations at hand to understand if feeding the optimization with additional characteristics improves the policies overall. To this purpose, we comment on the graphs of Fig. 1.

The fact that the cumulative density functions do not shift to the right when going from the upper graphs to the lower graphs does not make a clear case for an overall superiority of triple characteristic policies. We illustrate this finding with a simple example: for the constrained double characteristic portfolio policy, the top two choices for $\gamma = 5$ are BTM–ERV and DIY–ERV with Sharpe ratios of 0.54 and 0.53, respectively (see Table 4). The constrained policy based on BTM, DIY and ERV has a Sharpe ratio of 0.51 for $\gamma = 5$. As such, an intuitive combination of three seemingly well performing characteristics does not necessarily add value. This is further confirmed by the ‘kitchen sink’ combination of MEQ, BTM, DIY, VAR, ERV and AGR (which have led to high Sharpe ratios in previous settings): for a risk aversion parameter of 5, it shows a Sharpe ratio of 0.51 with the leverage constraint and 0.4 without.

However, we document a stability in the indicators which are associated to high Sharpe ratios. With respect to unconstrained triple characteristic policies and among the 10 highest Sharpe ratios for $\gamma = 5$, the BTM, ERV and AGR characteristics are those with the highest number of occurrences (5 each). For constrained policies, it is ERV (seven times among the top 10 Sharpe ratios), MEQ (six times) and DIY (five times). This is consistent with what we observed for double characteristic portfolios, except that VAR (resp. MEQ) featured more (resp. less) within the best combinations. The momentum attribute is never present among the best performing combinations.

Our results show that the addition of more characteristics does not necessarily increase the performance of the strategy. For example, when comparing the classical MEQ–BTM pair with the MEQ–BTM–MOM triplet across constrained versus unconstrained optimizations with $\gamma = 1$ or $\gamma = 5$, we find that the triplet never outperforms the pair in terms of Sharpe ratio. Moreover, the triplet is always associated to higher transaction costs. Accordingly, in our

Table 6

Time-varying leverage and risk aversion. This table displays the results for all key indicators based on the double characteristic portfolio policy *BTM-MEQ* when leverage and risk aversion are driven by state variables. The equal mix of variables is the mean of all three variables, once they have been normalized so that their mean is equal to one. For the base case, we fixed $\gamma = 5$ and $\delta_T = N_T^{-1}$. A list of all abbreviations can be found in Appendix A. Transaction costs and HHI are expressed in percent.

State variable	Key indicators					
	Vol	SR	TC	PNW	CER	HHI
Volatility	0.197	0.523	1.56	0.132	0.046	0.49
Credit spread	0.190	0.578	1.37	0.119	0.062	0.52
Term spread	0.265	0.709	3.11	0.174	0.037	1.48
Equal mix	0.194	0.581	1.58	0.134	0.059	0.51
Constant leverage and γ	0.188	0.467	1.05	0.092	0.045	0.31

framework, an investor would be better off with only the size and book-to-market attributes. Thus, it seems that compiling characteristics does not improve performance, apart for a few particular cases which can only be identified via data mining. Overall, we believe that retaining two (or three) characteristics is a reasonable choice for investors.

5. Robustness checks and sensitivity analysis

The goal of this section is to show the robustness of our results to variations in the allocation parameters. We also depend the understanding of the methodology by performing several sensitivity analyses and emphasize the robustness of our results by showing that our conclusions hold when using CRRA utility.

5.1. Sensitivity to the leverage constraint

Investors have heterogeneous leverage constraints, not only among them, but also across time (Ang et al. (2011) show that between 2005 and 2010, average gross leverage for equity hedge funds oscillated between 1.1 and 1.3.). In all previous numerical applications so far, we have set the bindingness of the leverage constraint to $\delta_T = N_T^{-1}$ since it is a reasonable compromise between zero and the full constraint. However, this level might be too tight or too loose, depending on whether the investor is a hedge fund or a pension fund. Table 7 shows the sensitivities of the key indicators with respect to the bindingness of the leverage constraint (δ_T) based on the double characteristic portfolio policy BTM-MEQ. In order to investigate the impact of variations in this parameter, we report the key indicators when $\delta_T = N_T^{-\kappa}$ and $\kappa \in (0.0, 0.1, \dots, 2)$ so that N_T ranges between N_T^{-2} and 1, which are the two bounds of Lemma 2.1.¹⁹

Note that the figures in Table 7 do not exactly converge to those of the equally-weighted benchmark because for some stocks one of the attributes (MEQ or BTM) was not available which excluded them from the optimization. As expected, the transaction costs and the proportion of negative weights decrease as the constraint becomes tighter (i.e. when κ increases). This effect is especially pronounced for $\gamma = 1$.

In the case of low risk aversion ($\gamma = 1$), both the Sharpe ratio and the volatility strongly decrease with κ . For a more moderate risk aversion, the Sharpe ratio decreases, whereas the volatility remains nearly constant. For $\gamma = 1$, we find that for $\kappa = 1.3$, the policies are close to a long-only policy (less than 1% of short-sales) and the Sharpe ratio raises to 0.46, from 0.37 when $\kappa = 1.9$ (highly constrained program). Fine-tuning the constraint therefore allows

¹⁹ Note that we do not report results for $\kappa < 0.4$ because the values do not change below this threshold.

Table 7

Sensitivity to changes in the bindingness of the leverage constraint. This table displays the results for all key indicators based on the double characteristic portfolio policy BTM-MEQ when decreasing the intensity of the leverage constraint. We fixed $\delta_T = N_T^{-\kappa}$ for $\kappa \in (0.4, 0.5, \dots, 1.9)$. The case $\kappa = 0.4$ is equivalent to the unconstrained policy. A list of all abbreviations can be found in Appendix A. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark. Transaction costs and HHI are expressed in percent.

κ	$\gamma = 1$						$\gamma = 5$					
	Vol	SR	TC	PNW	CER	HHI	Vol	SR	TC	PNW	CER	HHI
1.9	0.18	0.37	0.67	0.00	0.10	0.18	0.18	0.35	0.65	0.00	0.03	0.18
1.8	0.18	0.37	0.69	0.00	0.10	0.18	0.18	0.36	0.66	0.00	0.03	0.18
1.7	0.19	0.38	0.71	0.00	0.10	0.18	0.18	0.36	0.67	0.00	0.03	0.18
1.6	0.19	0.39	0.73	0.00	0.10	0.18	0.18	0.37	0.69	0.00	0.03	0.18
1.5	0.19	0.41	0.77	0.00	0.11	0.18	0.19	0.37	0.71	0.00	0.03	0.18
1.4	0.19	0.43	0.82	0.00	0.11	0.19	0.19	0.39	0.74	0.00	0.03	0.19
1.3	0.20	0.46	0.89	0.01	0.12	0.20	0.19	0.40	0.79	0.01	0.03	0.20
1.2	0.20	0.50	1.00	0.03	0.13	0.22	0.19	0.42	0.85	0.02	0.04	0.22
1.1	0.21	0.55	1.14	0.07	0.14	0.27	0.19	0.44	0.94	0.05	0.04	0.26
1.0	0.22	0.61	1.35	0.13	0.15	0.35	0.19	0.47	1.05	0.09	0.05	0.31
0.9	0.23	0.68	1.62	0.19	0.17	0.48	0.19	0.50	1.21	0.15	0.05	0.39
0.8	0.25	0.76	2.01	0.24	0.19	0.74	0.19	0.51	1.38	0.20	0.05	0.50
0.7	0.27	0.87	2.55	0.30	0.23	1.21	0.18	0.52	1.58	0.22	0.05	0.66
0.6	0.30	0.96	3.24	0.33	0.27	2.03	0.19	0.53	1.68	0.24	0.05	0.79
0.5	0.35	1.05	4.20	0.36	0.32	3.60	0.19	0.53	1.68	0.24	0.05	0.79
0.4	0.38	1.11	5.12	0.38	0.35	5.27	0.19	0.53	1.68	0.24	0.05	0.79

us to keep very low leverage while improving the risk-adjusted potential of the portfolio.

Lastly, we see that when risk aversion is low, the CER increases when the constraint loosens since the investor does not mind to holding a highly volatile portfolio. However, this is no longer true for a moderate risk aversion and, consequently, no economic benefit is obtained for a constraint level which is above N_T^{-1} .

5.2. Sensitivity to sample size

From Eq. (4), we see that θ_T is chosen such that it would have maximized the expected utility of the investor given past values of characteristics and returns. In our base case computations, we have used rolling samples of $\tau = 10$ years to successively calibrate the values of θ_T . This choice is somewhat arbitrary and it is legitimate to wonder whether shorter sample sizes would lead to improvements or not.²⁰ Indeed, large samples give old data as much importance as recent data, while it is not obvious that the cross-sectional predictive power of characteristics remains stable over time. As such, calibrating on smaller samples may allow θ_T to adjust more rapidly, especially in times of turbulence.

Table 8 reports the impact of estimation sample size on all indicators for the double characteristic portfolio policy BTM-MEQ. For unconstrained portfolios (first five columns), we observe that the Sharpe ratio is slightly decreasing with τ , but we cannot observe any monotonous impact on volatility, nor on CER. In contrast, transaction costs and leverage decrease with τ . For transaction costs, this is quite straightforward, as longer sample size implies more stability of θ_T through time.

When leverage constraints are enforced (last six columns), the Sharpe ratio decreases with τ , while the volatility remains constant, which means that average returns are higher when the estimation sample size is smaller. Transaction costs and leverage are strongly reduced compared to the unconstrained policies. In short, this means all indicators remain stable when reducing the estimation sample with the Sharpe ratio increasing marginally. This is possible because of the regularization effect of the constraint. Without the constraint, reducing the estimation sample implies a bad conditioning of the sample covariance matrix of the

characteristics-based portfolios and θ_T becomes progressively degenerate. For example, when $\tau = 2$, Σ_T in (6) is singular and it is possible to compute θ_T using the Moore–Penrose inverse. This leads to a negative Sharpe ratio.

Overall, the unconstrained optimization can only generate reasonable weights if the sample size is large enough. This drawback is circumvented when adding the constraint. Moreover, constrained policies seem to deliver better risk-adjusted performance with shorter sample sizes.

For the sake of completeness, we also have computed the indicators across all sample sizes, all risk aversions and all characteristics. Over all of these combinations, switching from unconstrained to constrained optimization leads to a reduction in transaction costs which is at least equivalent to 10% and can reach up to 93%. For the proportion of negative weights, the minimum reduction is 37% and the maximum one is 90%. The impact on volatility oscillates between a 20% rise (a rare phenomenon which only appears for $\gamma = 10$) and a decrease of 71%. In line with our prior results, there is no particular pattern for the evolution of the Sharpe ratio.²¹

5.3. Impact of transaction costs on the Sharpe ratio

To this point, we have reported the gross Sharpe ratio in all tests without the inclusion of transaction costs. However, as suggested by the recent literature, transaction costs can have a significant impact on characteristics-based investment strategies. For example, Novy-Marx and Velikov (2016) find that the mitigation of transaction costs is crucial when looking at net returns of such strategies. So far, our findings have underlined that the presence of the constraint shrinks both the turnover and the transaction costs. Hence, when considering the impact of the constraint, we expect an even more favorable impact on transaction cost-adjusted Sharpe ratios. Table 9 shows the results of three different characteristics-based strategies when comparing raw- and transaction cost adjusted Sharpe ratios.²² The transaction cost-adjusted Sharpe ratio is computed as the first moment of returns minus the risk free rate minus the transaction costs divided by the second moment of returns (transaction costs are subtracted from returns

²⁰ Brandt et al. (2009) also consider 10 years of past data. In our mean–variance framework, we recall that τ must be strictly larger than the number of characteristics for unconstrained policies to be well defined.

²¹ The results are available upon request.

²² The kitchen sink strategy contains the best single and double firm characteristics: MEQ, BTM, DIY, VAR, and AGR.

Table 8

Sensitivity to the estimation sample size. This table displays the results for all key indicators for policies based on the double characteristic portfolio policy *BTM-MEQ* when the estimation sample size (τ) decreases. In all cases, the allocation process starts in 1980: for $\tau < 10$, the data is truncated accordingly. In the case of the constrained policies, we fixed $\delta_T = N_T^{-1}$ and $\gamma = 5$. A list of all abbreviations can be found in [Appendix A](#). Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark. Transaction costs and HHI are expressed in percent.

τ	Unconstrained policies						Constrained policies					
	Vol	SR	TC	PNW	CER	HHI	Vol	SR	TC	PNW	CER	HHI
10	0.17	0.61	2.40	0.28	0.06	1.47	0.19	0.47	1.05	0.09	0.04	0.31
9	0.19	0.62	2.39	0.28	0.06	1.28	0.19	0.49	1.06	0.09	0.05	0.28
8	0.19	0.61	2.57	0.26	0.06	1.25	0.19	0.50	1.08	0.08	0.05	0.25
7	0.19	0.60	2.61	0.25	0.05	1.24	0.19	0.49	1.12	0.08	0.05	0.23
6	0.20	0.61	2.69	0.24	0.05	1.38	0.19	0.49	1.11	0.08	0.05	0.21
5	0.21	0.65	2.72	0.24	0.06	1.51	0.20	0.50	1.14	0.08	0.04	0.20
4	0.19	0.62	3.59	0.26	0.05	1.88	0.19	0.52	1.18	0.07	0.05	0.18
3	0.21	0.70	4.43	0.29	0.05	2.29	0.19	0.52	1.22	0.06	0.05	0.17

Table 9

Comparison of raw- and transaction cost-adjusted Sharpe ratios. This table displays the results for raw- and transaction cost-adjusted Sharpe ratios for the constrained and unconstrained characteristics-based portfolio policy using constraint (3). The *Kitchen sink* set contains the assortment of the best performing firm characteristics: MEQ, BTM, DIY, VAR, ERV, and AGR. We compute the results of the portfolio policy using different risk-aversion values (γ) of 1, 5 and 10. All results are based on $\delta_T = N_T^{-1}$. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally-weighted benchmark.

γ	Raw Sharpe ratio						Transaction cost-adjusted Sharpe ratio					
	Unconstrained			Constrained			Unconstrained			Constrained		
	1	5	10	1	5	10	1	5	10	1	5	10
MEQ-BTM	1.33	0.61	0.23	0.61	0.47	0.23	1.19	0.48	0.11	0.55	0.41	0.18
MEQ-BTM-MOM	1.33	0.54	0.14	0.62	0.47	0.24	1.13	0.31	−0.08	0.55	0.41	0.17
Kitchen sink	0.73	0.40	0.28	0.63	0.51	0.35	0.21	0.13	−0.19	0.57	0.46	0.27

at each date). As in the previous case, bold figures indicate a statistically higher Sharpe ratio at the 10% confidence level based on the [Ledoit and Wolf \(2008\)](#) test.

The Sharpe ratios are in line with our prior findings on raw Sharpe ratios, but tend to be even higher. Indeed, we observe that the constraint has a larger impact when controlling for transaction costs for many firm characteristics. This is particularly true for the kitchen sink combination: the transaction cost-adjusted Sharpe ratios of constrained portfolios are equal to three times their unconstrained counterparts. This increase in net Sharpe ratio is a direct consequence of the reduction of transaction costs induced by the constraint: the numerator of the unconstrained Sharpe ratio is strongly negatively impacted while the denominator remains roughly unchanged. When constraints are enforced, the variation of both terms is much smaller. From a practical point of view, this means that the implementation of constrained portfolio policies becomes even more beneficial when net returns are considered.

5.4. CRRA preferences

In this subsection, we explore the impact of the constraint in the setting of CRRA preferences, notwithstanding the technical caveats related to this choice of utility function ($u(x) = -x^{-\gamma}/\gamma$).

We gather our results in [Table 10](#) for all indicators when the intensity of the constraint varies and when assuming a CRRA utility function. As in [Table 7](#), we find a general trend: a loosening of the constraint leads to higher transaction costs, volatility, negative portfolio weights and Sharpe ratios. However, in contrast to the quadratic utility case, the absolute levels are significantly higher (few investors can cope with a volatility of 101%). An HHI of 50% means that the magnitude of the weights is far beyond what can be implemented in practice. Transaction costs go up to 17.3% for $\kappa = 0.1$, the proportion of negative weights up to 45% and the CER down to minus 138%. These large values do not only underline the necessity of our constraints but also the problems associated with CRRA discussed in [Section 2.2](#). Lastly, we observe that the highest values of the CER are reached for values of κ close to 1, which

Table 10

Sensitivity to leverage constraint with CRRA preferences. This table displays the results for all key indicators based on the double characteristic portfolio policy *BTM-MEQ* for CRRA when the intensity of the leverage constraint increases. We fixed $\gamma = 5$ and $\delta_T = N_T^{-\kappa}$ for $\kappa \in (0.1, 0.2, \dots, 1.6)$. A list of all abbreviations can be found in [Appendix A](#). Transaction costs and HHI are expressed in percents.

κ	Key indicators					
	Vol	SR	TC	PNW	CER	HHI
0.1	1.01	1.26	17.3	0.45	−1.38	52.72
0.2	0.77	1.22	12.2	0.44	−0.60	27.56
0.3	0.60	1.18	8.8	0.43	−0.22	14.73
0.4	0.47	1.12	6.4	0.41	−0.04	7.91
0.5	0.38	1.04	4.7	0.39	0.04	4.29
0.6	0.32	0.96	3.5	0.36	0.07	2.36
0.7	0.28	0.87	2.7	0.32	0.08	1.34
0.8	0.25	0.78	2.1	0.26	0.07	0.79
0.9	0.23	0.69	1.7	0.19	0.07	0.50
1.0	0.21	0.62	1.3	0.13	0.06	0.35
1.1	0.20	0.55	1.1	0.06	0.05	0.27
1.2	0.20	0.49	1.0	0.02	0.05	0.22
1.3	0.19	0.45	0.9	0.01	0.04	0.20
1.4	0.19	0.42	0.8	0.00	0.04	0.19
1.5	0.19	0.39	0.7	0.00	0.03	0.18
1.6	0.19	0.36	0.7	0.00	0.03	0.18

confirms that our initial choice of δ_T is relevant and that the policy should be constrained so as to generate reasonable (i.e. implementable) results.

5.5. Comparison with [Brandt et al. \(2009\)](#)

This subsection is intended to show the added value on investor welfare of a quadratic constraint compared to the approach of [Brandt et al. \(2009\)](#) who propose a heuristic way of obtaining long-only portfolios by setting negative weights to zero and rescaling positive weights so as to fulfill the budget constraint. As in [Fan et al. \(2012\)](#) and [Coqueret \(2015\)](#), we argue that a *small* proportion of negative weights is an effective compromise between large levels of leverage and no short-selling at all.

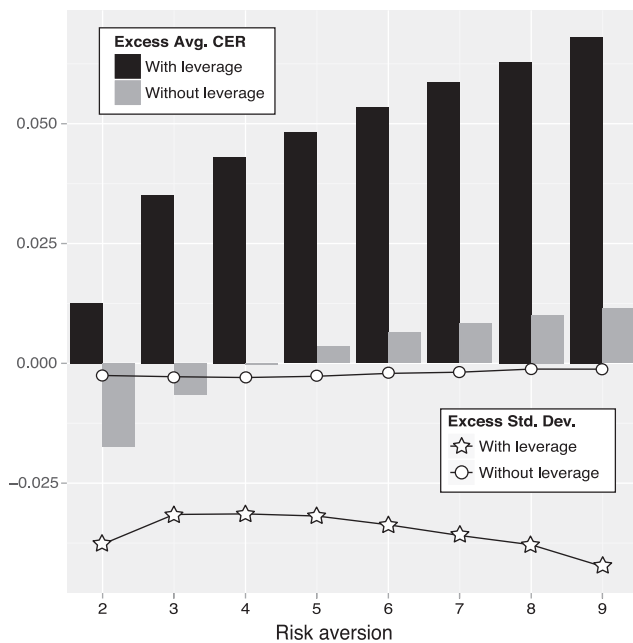


Fig. 3. Comparison with Brandt et al. (2009). The bars represent the excess average CER (over the 10 subsamples) of the constrained strategy, net of the average of either alternatives of Brandt et al. (2009) (unconstrained or long-only after setting negative weights to zero and rescaling). Likewise, Excess Std. Dev. refers to the excess standard deviation of the CER.

In order to compare our methodology with the two alternatives of Brandt et al. (2009), we adopt the following protocol. First, given the caveats of the power utility function shown in the previous subsection, we restrain our analysis to quadratic utility. Also, for the sake of robustness, we will report the results for several values of risk aversion. Brandt et al. (2009), carry out their optimization over three characteristics: firm size (MEQ), book-to-market ratio (BTM) and past annual return (MOM). Accordingly, we will proceed with the exact same triplet of characteristics. Lastly, we perform the allocation on ten subsamples extracted from our original database. This will allow us to have several points for comparison purposes and reduces the risk of sample-dependence of our results.²³ The subsets are constructed as follows: Given the original ordering of the companies in the dataset (in increasing Compustat GVKEY), we split the firms according to the last number of their ranking (firms n°6, 16, 26, 36, etc. belong to the same group). We then proceed with the allocation steps and compute the CER of all strategies. We choose the CER as the driving indicator because it synthesizes the first two moments of returns, risk aversion and transaction costs in one simple metric. We restrict the values of risk aversion to all integers between 2 and 9 because they embed all the levels that seem acceptable by asset managers on the market (values above 9 do not change the allocation much and those below 2 lead to very aggressive and volatile portfolios).

Fig. 3 shows the excess average in CER of the constrained strategy, net of the average of either alternatives of Brandt et al. (2009) (unconstrained or long-only after cutting negative weights and rescaling). The black bars represent the case in which leverage is allowed whereas the gray bars show the results when leverage is set to zero.²⁴ Further, we show the excess standard deviation in which a negative value stands for a lower return volatility when applying our constraint.

When leverage is allowed (black bars), we find a positive average excess CER for all levels of risk aversion. Moreover, the standard deviations of the CER show that this gain is quite stable across different investment universes. Notably, the gains are increasing with γ . When leverage is prohibited (gray bars), we find a positive impact of our constraint for levels of risk aversion of above $\gamma = 4$. For risk seeking investors it is more favorable to use the truncated version of Brandt et al. (2009) instead of setting the leverage to zero via the constraint. However, one should be aware of the fact that cutting off negative weights and rescaling afterwards creates large bets on fewer stocks. Consequently, this approach significantly reduces the diversification of the portfolio.²⁵

Overall, we find that when directly comparing our constraint to the framework of Brandt et al. (2009) the introduction of the constraint leads to an increase in investor welfare for nearly all cases. This lets us conclude that in order to fully exploit the possibilities and advantages of the original optimization, investors should fine-tune the leverage of their portfolios.

6. Conclusion

We address several so far unsolved shortcomings of characteristics-based portfolio optimization. With the introduction of a leverage constraint into the modified framework of Brandt et al. (2009), we present an approach which reduces the dispersion of weights around an agnostic prior: the equally-weighted portfolio. Based on an out-of-sample empirical analysis from 1969 to 2013, we find that portfolio policies which include our constraint exhibit significantly lower levels of leverage compared to unconstrained policies, while keeping up similar Sharpe ratios and CER. The constraint facilitates the implementation in practice and is attractive for investors who are restricted to given levels of leverage. It is also flexible and can be both investor-dependent and time-varying.

In addition, we observe that our constraint leads to a significant reduction in discrepancy across characteristics, which lowers the odds of abnormal underperformance (very low or negative Sharpe ratios) subsequent to a poor choice of characteristics. With respect to variations in risk aversion, it turns out that constrained policies have a smaller dispersion in Sharpe ratios when switching from high to low risk aversion. Consequently, constrained policies seem less model-dependent and more robust. This is typically useful because quantifying an investor's risk aversion is usually not straightforward in practice. We also find that constrained policies have much lower levels of turnover and transaction costs compared to their unconstrained counterparts. Finally, constraints are also advantageous because they imply exposures to characteristics (and the possibly related risk factors) that do not evolve too abruptly.

With regard to the selection of characteristics, we have worked with a set of 12 firm characteristics, thereby broadening the canonical size-value-momentum paradigm. In fact, while we acknowledge that market equity and book-to-market indicators do yield above average Sharpe ratios, we do not find that past returns (i.e. momentum) are likely to add any further value. This latter attribute is unstable and often generates higher turnover. According to our findings, firm characteristics which should be considered by investors include: dividend yields, variance of returns, variance of earnings, and asset growth. Lastly, we conclude that adding more than two characteristic does not benefit the policies.

Overall, our approach enhances the applicability of characteristics-based portfolio choice and provides new insights into numerical optimization with respect to risk aversion, sample size, and constraint tightness.

²³ We thank Abraham Lioui for suggesting this enhanced protocol.

²⁴ In our framework, the constraint is tightened until all negative weights disappear.

²⁵ In unreported tests we find significantly higher degrees of diversification when applying the constraint compared to truncating the negative weights.

Table 11

Abbreviations of firm characteristics, key indicators and input parameters. The table shows the definition of the used abbreviations for all firm characteristics, key indicators and input parameters within the article. The firm characteristics show the abbreviation of all 12 individual firm characteristics that were regarded in our analysis. A more detailed description can be found in [Appendix B](#). The group of key indicators consist of the measures which are used to evaluate the performance and characteristics of the constrained and unconstrained portfolio policies. A detailed description of all key indicators can be found in [Section 3.2](#). Finally, we specify input parameters as those variables which are varied for the understanding of the input sensitivities of the optimization algorithm. N_T is the number of firms within the scope of the optimization at time T .

Firm characteristics	
AGR	Year-over-year asset growth
BTM	Book-to-market
CFA	Cash-flow over assets
CUE	Annual change in earnings
DIY	Current dividend yield
ERV	Earnings volatility
GMV	Absolute annual variation in gross-margin
LEV	Leverage-ratio
MEQ	Market equity
MOM	Momentum return
ROA	Return on assets
VAR	Return variance
Key indicators	
PNW	Proportion of negative weights
SR	Sharpe ratio
TC (%)	Transaction costs (in percent)
Vol	Portfolio volatility
CER	TC-adjusted certainty equivalent return
HHI	Average Herfindhal-Hirschman Index
Model parameters	
γ	Risk aversion parameter
$\delta_T = N_T^{-k}$	Tightness of the leverage constraint
τ	Estimation sample size in years

Appendix A. Abbreviations

See [Table 11](#).

Appendix B. Data

We proceed in several steps. First, we restrict our sample to all companies of the North America Compustat database which have at least 10 years of business activity. Consequently, all firm characteristics and return data come from the CRSP/Compustat Merged Database (CCM). We calculate the firm characteristics at the end of June each year from 1964 to 2013 and the first years are kept to calculate lagged returns (the final firm characteristics are reported from 1967 onwards). We consider common/ordinary security types only ($tpci = 0$) to avoid influences of issue-specific attributes. We use annual data for fundamentals and monthly data for prices and total returns.

In the following we describe the calculation of our firm characteristics. As in [Brandt et al. \(2009\)](#), we calculate the company's book equity as total assets (CCM item: AT) minus total liabilities (CCM item: LT) plus deferred taxes and investment tax credit (CCM item: TXDITC) minus the preferred stock value (CCM item: PSTK). A company's market equity (MEQ) is determined as the price per share (CCM item: PRCCM) times the number of common shares outstanding (CCM item: CSO). The book-to-market ratio (BTM) is defined as book equity divided by market equity. We calculate the current dividend yield DIY as total dividends (CCM item: DVP and DVC) divided by the number of common shares outstanding times the share price. The computation of leverage (LEV) follows [Bhandari \(1988\)](#): the leverage of a company is the difference of total assets and the book value of equity divided by market equity. Momentum (MOM) is based on returns from $t-12$ months to $t-2$ months and relies on price data only. In contrast, we evaluate

the variance of the returns (VAR) based on total returns over the past 60 months. Return on assets (ROA) is seen as the ratio of income before extraordinary items (CCM item: IB) and total assets. The firm's cash-flow over assets (CFA) is calculated as net income (CCM item: NI) plus depreciation (CCM item: DPC) minus the change in net working capital (CCM item: WCAPCH) minus capital expenditures (CCM item: CAPX) divided by the firm's total assets. GMV is the five year absolute variation in the firm's gross margin, whereas the margin is calculated as revenue (CCM item: REVT) minus costs of goods sold (CCM item: COGS) divided by total sales. Finally, we introduce earnings volatility (ERV) as the standard deviation of the firm's return-on-assets over the past 20 quarters. Both the annual change in earnings (CUE) and asset growth (AGR) are computed as in [Hand and Green \(2011\)](#): simple growth rate of the change in net income or total assets, respectively.

Lastly, we exclude negative data values for MEQ, BTM, DIY, LEV and VAR. We also eliminate the 20% of the firms with the smallest market equity and all values which lie five standard deviations above (or below) the cross-sectional average each year.

Appendix C. Proof of Proposition 2.1

We recall the following matrix notations: \mathbf{w}_T and \mathbf{r}_{t+1} are the $(N \times 1)$ vectors corresponding to $w_{i,T}$ and $r_{i,t+1}$ respectively and \mathbf{x}_t is the $(N \times F)$ concatenation of the $\mathbf{x}_{i,t}$ vectors. The Lagrangian associated to the problem (4) is

$$\begin{aligned}
 G(\boldsymbol{\theta}_T) &= \frac{1}{T} \sum_{t=T-\tau}^{T-1} \sum_{i=1}^{N_T} (\bar{\mathbf{w}}_{i,t} + \boldsymbol{\theta}'_T \mathbf{x}_{i,t}) r_{i,t+1} \\
 &\quad - \frac{\gamma}{2T} \sum_{t=T-\tau}^{T-1} \left[\sum_{i=1}^{N_T} (\bar{\mathbf{w}}_{i,t} + \boldsymbol{\theta}'_T \mathbf{x}_{i,t}) r_{i,t+1} \right]^2 - \lambda (\boldsymbol{\theta}'_T \mathbf{x}'_T \mathbf{x}_T \boldsymbol{\theta}_T - \delta_T) \\
 &= \frac{1}{T} \sum_{t=T-\tau}^{T-1} (\mathbf{w}'_t + \boldsymbol{\theta}'_T \mathbf{x}'_t) \mathbf{r}_{t+1} \\
 &\quad - \frac{\gamma}{2T} \sum_{t=T-\tau}^{T-1} (\bar{\mathbf{w}}'_t + \boldsymbol{\theta}'_T \mathbf{x}'_t) \mathbf{r}_{t+1} \mathbf{r}'_{t+1} (\mathbf{x}_t \boldsymbol{\theta}_T + \bar{\mathbf{w}}_t) \\
 &\quad - \lambda (\boldsymbol{\theta}'_T \mathbf{x}'_T \mathbf{x}_T \boldsymbol{\theta}_T - \delta_T)
 \end{aligned}$$

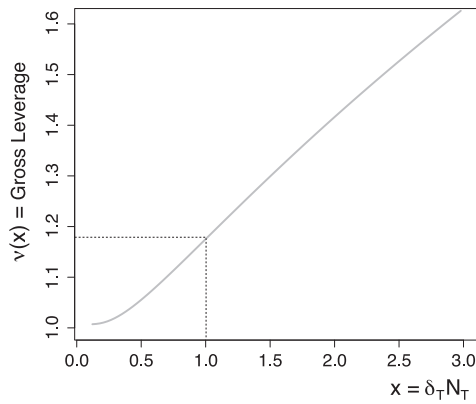
and hence,

$$\begin{aligned}
 \frac{\partial G}{\partial \boldsymbol{\theta}_T}(\boldsymbol{\theta}_T) &= \frac{1}{T} \sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} \\
 &\quad - \frac{\gamma}{T} \sum_{t=T-\tau}^{T-1} (\mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \bar{\mathbf{w}}_t + \mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \mathbf{x}_t \boldsymbol{\theta}_T) - 2\lambda \mathbf{x}'_T \mathbf{x}_T \boldsymbol{\theta}_T,
 \end{aligned}$$

so that the first order condition implies that

$$\begin{aligned}
 \boldsymbol{\theta}_T &= \left[2\lambda T \mathbf{x}'_T \mathbf{x}_T + \gamma \sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \mathbf{x}_t \right]^{-1} \\
 &\quad \times \left[\sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} - \gamma \sum_{t=T-\tau}^{T-1} \mathbf{x}'_t \mathbf{r}_{t+1} \mathbf{r}'_{t+1} \bar{\mathbf{w}}_t \right].
 \end{aligned}$$

We underline that the conditions $\tau > F_T$ and $N_T > F_T$ ensure that the inverse matrix is well-defined. The remaining degree of freedom, λ , is chosen such that condition (3) is satisfied. The second order condition straightforwardly implies that the solution is indeed a maximum point. If δ_T is very large, then the problem is unconstrained. If the problem is indeed constrained, then as λ increases to infinity, $\boldsymbol{\theta}'_T \mathbf{x}'_T \mathbf{x}_T \boldsymbol{\theta}_T$ continuously (but not necessarily monotonously) decrease to zero and any value δ_T can be reached (this can be formally shown using the strictly positive (since $N_T > F_T$) eigenvalues of $\mathbf{x}'_T \mathbf{x}_T$, as in [A in Coqueret \(2015\)](#)).

Fig. D.4. Plot of v .

Appendix D. Proof of Lemma 2.1

We have

$$\mathbb{E}[L_T] = \mathbb{E}\left[\sum_{i=1}^{N_T} |w_{iT}|\right] = \mathbb{E}\left[\sum_{i=1}^{N_T} \left|\frac{1}{N_T} + y_{iT}\right|\right] = \sum_{i=1}^{N_T} \mathbb{E}\left[\left|\frac{1}{N_T} + y_{iT}\right|\right]$$

Now, the $|1/N_T + y_{iT}|$ are i.i.d. folded Normal variables. Using the formula for the mean of the folded normal distribution (see Chakraborty and Chatterjee (2013) for instance), we get

$$\mathbb{E}[L_T] = \sum_{i=1}^{N_T} \left[\sqrt{\frac{\delta_T}{N_T}} \times \sqrt{\frac{2}{\pi}} e^{-(2\delta_T N_T)^{-1}} + \frac{1}{N_T} (1 - 2\Phi(-(\delta_T N_T)^{-1/2})) \right],$$

which, after elementary simplifications, yields the result, where

$$v(x) = 1 + \sqrt{2x/\pi} e^{-1/(2x)} - 2\Phi(-x^{-1/2}).$$

We plot this function below. Moreover, this function is strictly increasing on \mathbb{R}_+ because:

$$v'(x) = \frac{(1+x)e^{-1/(2x)}}{\sqrt{2\pi}x^{3/2}} - \frac{e^{-1/(2x)}}{\sqrt{2\pi}x^{3/2}} = \frac{xe^{-1/(2x)}}{\sqrt{2\pi}x^{3/2}} > 0, \quad x > 0.$$

References

- Ang, A., Gorovyy, S., Van Inwegen, G.B., 2011. Hedge fund leverage. *Journal of Financial Economics* 102 (1), 102–126.
- Arnett, R.D., Hsu, J., Moore, P., 2005. Fundamental indexation. *Financial Analysts Journal* 61 (2), 83–99.
- Asness, C., Frazzini, A., Pedersen, L.H., 2013. Quality minus junk. Working paper available at SSRN.
- Bhandari, L.C., 1988. Debt/equity ratio and expected common stock returns: empirical evidence. *Journal of Finance* 43, 507–528.
- Boudt, K., Wauters, M., Ardia, D., 2014. Characteristic-based equity portfolios: economic value and dynamic style-allocation, working paper available at SSRN.
- Brandt, M.W., Santa-Clara, P., Valkanov, R., 2009. Parametric portfolio policies: exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies* 22 (9), 3411–3447.
- Brown, D.B., Smith, J.E., 2011. Dynamic portfolio optimization with transaction costs: heuristics and dual bounds. *Management Science* 57 (10), 1752–1770.
- Campbell, J., Lo, A., MacKinlay, A., 1997. *The Econometrics of Financial Markets*. Princeton University Press.
- Carhart, M.M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52 (1), 57–82.
- Chakraborty, A.K., Chatterjee, M., 2013. On multivariate folded normal distribution. *Sankhya: The Indian Journal of Statistics* 75 (1), 1–15.
- Coqueret, G., 2015. Diversified minimum variance portfolios. *Annals of Finance* 11 (2), 221–241.
- Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance* 52 (1), 1–33.
- Daniel, K., Titman, S., 1998. Characteristics or covariances? *The Journal of Portfolio Management* 24 (4), 24–33.
- Demchuk, A., Gibson, R., 2006. Stock market performance and the term structure of credit spreads. *Journal of Financial and Quantitative Analysis* 41 (04), 863–887.
- DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R., 2009a. A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Management Science* 55 (5), 798–812.
- DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22 (5), 1915–1953.
- Estrella, A., Mishkin, F.S., 1998. Predicting us recessions: financial variables as leading indicators. *Review of Economics and Statistics* 80 (1), 45–61.
- Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427–465.
- Fan, J., Zhang, J., Yu, K., 2012. Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association* 107 (498), 592–606.
- Farnsworth, G., 2014. Strategic hedge fund leverage and investor welfare: A holdings-based approach. Working Paper.
- Garleanu, N., Pedersen, L.H., 2011. Margin-based asset pricing and deviations from the law of one price. *Review of Financial Studies* 26 (6), 1980–2022.
- Garleanu, N., Pedersen, L.H., 2013. Dynamic trading with predictable returns and transaction costs. *Journal of Finance* 68, 2309–2340.
- Geweke, J., 2001. A note on some limitations of CRRA utility. *Economics Letters* 71, 341–345.
- Goetzmann, W.N., Kumar, A., 2008. Equity portfolio diversification. *Review of Finance* 12 (3), 433–463.
- Goto, S., Xu, Y., 2015. Improving mean variance optimization through sparse hedging restrictions. *Journal of Financial and Quantitative Analysis* 50 (6), 1415–1441.
- Hand, J.R., Green, J., 2011. The importance of accounting information in portfolio optimization. *Journal of Accounting, Auditing & Finance* 26 (1), 1–34.
- Hentschel, L., Long, J.B., 2004. Numerical solution of the static portfolio problem for power utility investors. Working paper.
- Hjalmarsson, E., Manchev, P., 2012. Characteristic-based mean-variance portfolio choice. *Journal of Banking & Finance* 36 (5), 1392–1401.
- Jacobs, B., Levy, K., 2013. Leverage aversion, efficient frontiers, and the efficient region. *Journal of Portfolio Management* 39 (3), 54–64.
- Jacobs, B., Levy, K., 2014. Traditional optimization is not optimal for leverage-averse investors. *Journal of Portfolio Management* 40 (2), 30–40.
- Jacobs, B.L., Levy, K.N., 2007. 20 myths about enhanced active 120-20 strategies. *Financial Analysts Journal* 63 (4), 19–26.
- Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. *Journal of Finance* 58 (4), 1651–1684.
- King, R., Plosser, C.I., Rebelo, S.T., 2002. Production, growth and business cycles: technical appendix. *Computational Economics* 20, 87–116.
- Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15 (5), 850–859.
- Lo, A.W., Patel, P.N., 2008. 130/30: The new long-only. *The Journal of Portfolio Management* 34 (2), 12–38.
- Markowitz, H., 1952. Portfolio selection. *Journal of Finance* 7 (1), 77–91.
- Novy-Marx, R., Velikov, M., 2016. A taxonomy of anomalies and their trading costs. *Review of Financial Studies* 29 (1), 104–147.
- Pflug, G.C., Pichler, A., Wozabal, D., 2012. The 1/N investment strategy is optimal under high model ambiguity. *Journal of Banking & Finance* 36 (2), 410–417.
- Plyakha, Y., Uppal, R., Vilkov, G., 2012. Why does an equal-weighted portfolio outperform value- and price-weighted portfolios? Working paper available at SSRN.
- Rytchkov, O., 2014. Asset pricing with dynamic margin constraints. *The Journal of Finance* 69 (1), 405–452.
- Walkshäusl, C., Lobe, S., 2010. Fundamental indexing around the world. *Review of Financial Economics* 19 (3), 117–127.
- Yoon, G., 2004. On the existence of expected utility with CRRA under STUR. *Economics Letters* 83 (2), 219–224.