Characteristics-Based Portfolio Choice with Short-Sale Constraints

Manuel Ammann, Guillaume Coqueret and Jan-Philip Schade*

January 13, 2015

Abstract

We show that characteristics-based portfolio choice requires a short-sale constraint for reasonable levels of leverage. In addition to the introduction of our new constraint we include 12 characteristics to our study, thereby extending the classical size, book-to-market and momentum paradigm. We discuss the sensitivity of key indicators to the choice of characteristics, to risk aversion and to estimation sample size showing that constrained policies are much less responsive to these parameters than their unconstrained counterparts. Finally, in the case of quadratic utility, we derive a semi-closed analytical form for the portfolio weights. Overall, we find that the constraint effectively reduces negative weights, decreases both volatility and transaction costs of the portfolios and decreases the risk of model misspecification.

JEL classification: C61, G11.

1 Introduction

The topic of characteristic-based investing is classical in Modern Portfolio Theory and gained popularity after Fama and French (1992) showed that the size and book-to-market ratio of companies are strong drivers of the cross-sectional differences in future returns. In contrast to the intuitive way of gaining factor exposure by selecting stocks with the desired attributes and weighting them equally or proportionally to value (i.e. market capitalization) or other accounting quantities, new methods have recently blossomed. 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 approaches based on optimization procedures.[1] Our article introduces a new short-sale[2] constraint into these optimization schemes which significantly reduces the policy's negative weights and, hence, makes it applicable for a broad group of investors.

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 overall dimensionality that is often problematic when the investment universe is large. For instance, a classical minimum variance optimization for the

*Guillaume Coqueret is with EDHEC Business School at Nice, France. Manuel Ammann and Jan-Philip Schade are with University of St. Gallen, Switzerland. We are grateful to Jens Jackwerth and seminar participants at the Finance Seminar in Constance and St. Gallen for helpful comments and suggestions.

[1]See for instance Brandt et al. (2009), Hand and Green (2011), Hjalmarsson and Manchev (2012) and Boudt et al. (2014).

[2]In the following we will use the terms 'short-selling' and 'leverage' interchangeably.

S&P 500 universe requires the computation of more than 125,000 covariances. Furthermore, by basing their investment decisions on risk factors, investors can transfer their believes in different return drivers directly into their portfolio weights. Consequently, such an approach enables each investor to create investment strategies according to his taste. An intuitive framework for such an implementation was proposed by Brandt et al. (2009) who suggest to model optimal portfolio weights by deviating from an initial benchmark using a linear function of normalized characteristics. This is achieved by optimizing the expected utility of the future wealth with respect to the loadings of the characteristics. However, we find that this approach has several so far unsolved caveats.

First, similarly to the classical Markowitz (1952) mean-variance portfolios, the characteristics-based mean-variance portfolios are usually very leveraged: the optimal solutions imply large negative weights and many stocks must be shorted. For example Brandt et al. (2009) report more than 40% of negative weights on average in their empirical results. In practice, such levels of leverage are unrealistic, especially because many investors have long-only policies. A possible solution, advocated by Brandt et al. (2009) is to set negative positions to zero and rescale the weights. Given that this approach truncates the investable set by approximately 40%, it seems sub-optimal. Apart of the high levels of leverage, unconstrained portfolios are known to generate notoriously high instability, asset turnovers, and large transaction costs (for instance, Brandt et al. (2009) display turnovers above 100%.). Second, the choice of firm characteristics used in the allocation process is essential. When they are processed into metrics computed at the firm level, they are usually numerous. For example, Arnott et al. (2005) use accounting figures such as book value, cash flows, revenue, sales and dividends. Asness et al. (2013) compute z-scores based on 21 characteristics which can be categorized in four groups: profitability, growth, safety and payout. However, not every firm characteristic will end in superior results. Our results show that when the choice of the underlying characteristics is poor, unconstrained policies display disappointing results, with unexpectedly low (or even negative) Sharpe ratios. Our third point of criticism relates to the CRRA utility functions used by Brandt et al. (2009). The advantage of this class of functions is that it takes the higher moments of the portfolio returns into account. In contrast, the main drawback is that the numerical optimization remains intractable. This can be especially tenuous given that the optimization is based on several interacting factors such as the individual risk aversion or the sample size. For example, we find a significant change within the optimization outcome when risk aversion fluctuates. This underlines that it is crucial to understand the impact of the different parameters of the optimization.

This article has two major goals. The first is to introduce a comprehensible and easy-to-implement extension to characteristics-based portfolio optimization which overcomes the aforementioned limitations. Our model is close to the one of Brandt et al. (2009), though, as in Hjalmarsson and Manchev (2012), we consider quadratic utility functions which allow for analytical solutions that are directly interpretable.[3] Similarly to Jagannathan and Ma (2003)

[3]If returns are assumed to be Gaussian, then the merit of CRRA functions vanishes. In this paper, because of the low

and DeMiguel et al. (2009), we argue that adding a simple constraint in the optimization scheme will often contribute to a strong reduction in the risk of characteristics-based portfolios. This novel constraint has several supplementary virtues. First, it entails reduced levels of leverage. This is notably attractive for investors who are sensitive to margin requirements.[4] Second, it lessens the sensitivity of the performance to the selection of firm characteristics and to risk aversion. Lastly, constrained policies are more stable in time, compared to unconstrained portfolios. Consequently, asset rotation is strongly reduced in the presence of the constraint and the corresponding transaction costs are mechanically curtailed. Given that our findings hold under several robustness tests, these improvements highlight the benefits an investor can expect from the methodology we propose.

Our second goal is to understand the performance of different firm characteristics within the optimization framework and the sensitivity of the results towards the implemented input parameters. The large amount of possible firm characteristics stand in contrast to the handful of attributes that are considered as inputs in the optimization schemes of Brandt et al. (2009) and Hjalmarsson and Manchev (2012) or to very specific combinations of accounting figures (e.g. Hand and Green (2011)). Brandt et al. (2009) argue that since market capitalization, book-to-market and past returns suffice to explain the cross-section of returns (as shown by Fama and French (1992) and Carhart (1997)), they are sufficiently good enough candidates to be fed 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 makes room for a broader set of characteristics consisting of both 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. Finally, we study the sensitivity of our results to variations in important input parameters such as risk aversion, estimation sample size and bindingness of the leverage constraint.

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 is devoted to extensions and robustness checks related to the optimal number of characteristics that should be considered, to the sensitivity of our results to sample size, the bindingness of the constraint and to the factor exposure of the portfolios. Finally, in Section 5, we conclude.

---

frequency at which accounting figures are released, we will 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 will tend to be closer to the normal distribution than short-horizon returns*". In this context, the benefits of CRRA functions are unclear.

[4]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 equal to 50%). The impact on prices of heterogeneous margin requirements across assets was studied by Garleanu and Pedersen (2011). Rytchkov (2014) considers a similar problem with only one asset which is subject to time-varying margin constraints. The idea of adding leverage constraints to portfolio choice optimization is not new (e.g. Grossman and Vila (1992)), but it is becoming increasingly popular (Jacobs and Levy (2013), Jacobs and Levy (2014)).

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 form:[5]

$$\boldsymbol{w}_T = \bar{\boldsymbol{w}}_T + \boldsymbol{x}_T \boldsymbol{\theta}_T, \tag{1}$$

where $\bar{\boldsymbol{w}}_T$ 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 $\boldsymbol{x}_T$ comprises the firm's characteristics normalized so that they have zero mean and unit variance. N_T will henceforth denote 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 will be updated at each rebalancing period. Accordingly, we seek to solve the following max-utility problem:

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

where $\boldsymbol{r}_{T+1}$ 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 $\boldsymbol{x}_T \boldsymbol{\theta}_T$ 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) appear. 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 to perform the maximization program (2) under the constraint

$$\boldsymbol{\theta}_T' \boldsymbol{x}_T' \boldsymbol{x}_T \boldsymbol{\theta}_T = \delta_T, \tag{3}$$

where δ_T satisfies the inequalities $(\boldsymbol{\theta}_T^0)' \boldsymbol{x}_T' \boldsymbol{x}_T \boldsymbol{\theta}_T^0 > \delta_T > 0$, 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 $\boldsymbol{x}_T \boldsymbol{\theta}_T$ 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 will reduce the magnitude

[5] 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.

of the elements of $\boldsymbol{x}_T\boldsymbol{\theta}_T$ which will increase the relative importance of the prior benchmark $\bar{\boldsymbol{w}}_T$. 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 will set the benchmark starting point to be the equally-weighted portfolio: $\bar{\boldsymbol{w}}_T = \mathbf{1}_{N_T}/N_T$, 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. (2009) 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: $\boldsymbol{\theta}_T'\boldsymbol{x}_T'\boldsymbol{x}_T\boldsymbol{\theta}_T = \delta_T \Leftrightarrow \boldsymbol{w}_T'\boldsymbol{w}_T = \delta_T + N_T^{-1}$ because the characteristics' matrix is normalized and the elements of $\boldsymbol{x}_T\boldsymbol{\theta}_T$ sum to zero. This is a convenient property because the constraint on the perturbations $\boldsymbol{x}_T\boldsymbol{\theta}_T$ 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 The derivation of $\boldsymbol{\theta}_T$ and its interpretations

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

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

where $t = T$ is the present date and $t = T - \tau$ is the first date of the estimation sample. For a vector $\boldsymbol{x}$, we write x_i its i^{th} element and for a matrix $\boldsymbol{X}$, $\boldsymbol{X}_i$ 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 $\boldsymbol{x}_T$ ensure that it will be the case. For implementation purposes, at a given date, the sample has a constant number of stocks (N_T) over all 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 $N_T > F_T$.[6]

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

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*

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

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

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

$$\boldsymbol{\Sigma}_T^{(P)} = \frac{1}{T} \sum_{t=T-\tau}^{T-1} \boldsymbol{x}_t' \boldsymbol{r}_{t+1} \boldsymbol{r}_{t+1}' \boldsymbol{x}_t. \tag{6}$$

This representation recalls those based on regressions of Hjalmarsson and Manchev (2012), except that new terms appear because of the constraint and the benchmark portfolio. The terms in (5) can be interpreted in the following way. First, the vectors $\boldsymbol{x}_t' \boldsymbol{r}_{t+1}$ correspond to the $(F_T \times 1)$ returns of portfolios with weights $\boldsymbol{x}_{iT}$ $(i = 1, \ldots, F_T)$ 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 $\boldsymbol{\sigma}_T^{(\bar{\boldsymbol{w}})}$ 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 (5) can be decomposed in two components: the first is equal to the maximum Sharpe portfolio[7] where the assets are the characteristics-based portfolios, and the second is an adjustment stemming from the covariance with the benchmark starting point.

Next, we discuss the mechanics of the optimization embedded in the formula (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 $\delta_T \downarrow 0$, $\lambda^* \to \infty$ and of course $\boldsymbol{\theta}_T \to \mathbf{0}$. More technically, the constraint (3) acts like a regularization à la Tikhonov[8]: if the sample size is too small ($\tau \leq F_T$), the unconstrained problem is ill-conditioned, but adding the constraint will guarantee 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).

The second important variable in (5) is the risk aversion parameter. In the second factor

[7] 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 $\left\{ \underset{\boldsymbol{w}}{\text{argmax}} \frac{\boldsymbol{w}'\boldsymbol{\mu}}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}},\ \mathbf{1}_N'\boldsymbol{w} = 1 \right\} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}_N \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$.

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

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

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

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

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

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 will be 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. But with the introduction of the constraint (3), the weights are progressively shrunk towards the benchmark portfolio as the constraint becomes tighter. Accordingly, setting a strong constraint is only efficient if the benchmark is well chosen.

2.3 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 will determine to what extent the final weights can differ from those of the equally-weighted starting point. We want to determine a non-parametric method in order to set a threshold which will generate 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 reasonable levels. We note y_{iT} for the elements of the vector $\boldsymbol{x}_T \boldsymbol{\theta}_T$ and hence the constraint (3) reads

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

The distribution of the y_i 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 $\boldsymbol{\theta}_T$. Given that we want to define a criterion which does not depend on any specific case, we adopt a non-parametric approach.

If the benchmark portfolio is equally weighted, the weights of Equation (1) will become negative whenever $\min(\boldsymbol{x}_T \boldsymbol{\theta}_T) < -1/N_T$. Moreover, the turnover between time T and time $T+1$ will be mostly generated by differences between $\boldsymbol{x}_T \boldsymbol{\theta}_T$ and $\boldsymbol{x}_{T+1} \boldsymbol{\theta}_{T+1}$ (rather than by the evolution of the weights between the rebalancing dates of the equally weighted benchmark). Consequently, both the proportion of negative weights and the turnover can be reduced by imposing that $\max(|\boldsymbol{x}_T \boldsymbol{\theta}_T|)$ be smaller than a given quantity.

The main indicator we are interested in is the proportion of negative weights of the portfolio policies (i.e. leverage). When the optimization is unconstrained, this proportion can be larger than 40% (Brandt et al. (2009)). In the limiting case $\delta_T = 0$, the weights are those of the

$1/N$ benchmark which are naturally positive. In the case of infinite risk aversion (minimum variance portfolios), Fan et al. (2012) and Coqueret (2014) show that it is worthwhile to allow for a small proportion of negative weights. Since the benchmark is equally weighted, we recall that there will be no short selling as long as $\min_i y_{iT} > -1/N_T$. We determine the level of constraint based on the following result.

Lemma 2.1. *If equalities (7) and $\sum_{i=1}^{N_T} y_{iT} = 0$ hold, then the smallest minimum value that can be reached by y_{iT} is $-\sqrt{(N_T - 1)\delta_T / N_T}$ and the largest minimum value is $-\sqrt{\delta_T / (N_T(N_T - 1))}$.*

The previous lemma shows that the smallest possible minimum weight is equal to $1/N_T - \sqrt{(N_T - 1)\delta_T / N_T}$ while the largest possible minimum weight is $1/N_T - \sqrt{\delta_T / (N_T(N_T - 1))}$. In order to avoid short sales, we would want these quantities to be positive: consequently, the range for δ_T should be between $(N_T(N_T - 1))^{-1}$ and $(N_T - 1)/N_T$. Asymptotically, this means between N_T^{-2} and 1. The first value is much to stringent: it would ensure zero leverage for sure. The second value is too permissive: it would ensure positive leverage almost surely. For very large investment universes, there is a large difference between these two bounds. In order to keep our empirical results as tangible as possible we want to retain one simple criterion for δ_T for the entire analysis. For this reason we consider an intermediate situation and choose the geometric mean of the extreme situations: $\delta_T = N_T^{-1}$. As we will show in the subsequent sections, this choice will generate no more than 10% of negative weights in nearly all the cases we have tested. Finally, we show in our robustness tests that the presented results also hold for other values of δ_T.

3 Empirical analysis

3.1 Data

The construction of the research universe is detailed in Appendix A. Figure 1 shows the cross-sectional average of each firm characteristic over the entire sample. From this figure, we see that the firm characteristics evolve differently over time. This is of major importance given that we want to benefit from the complementarity of these firm characteristics. In total the number of companies within our sample is growing over time from 1,353 (in 1969) to a maximum of 2,652 (in 2003).

In addition to the sole time-series development of our firm characteristics, Table 1 provides the average correlation between all reported characteristics. Together, both illustrations allow to verify that all retained characteristics carry non-redundant information.[9] The highest observed correlation is 52%, between the cash-flow over assets (*CFA*) and the return on assets (*ROA*).

[9] We have also computed 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 were highly correlated with 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.

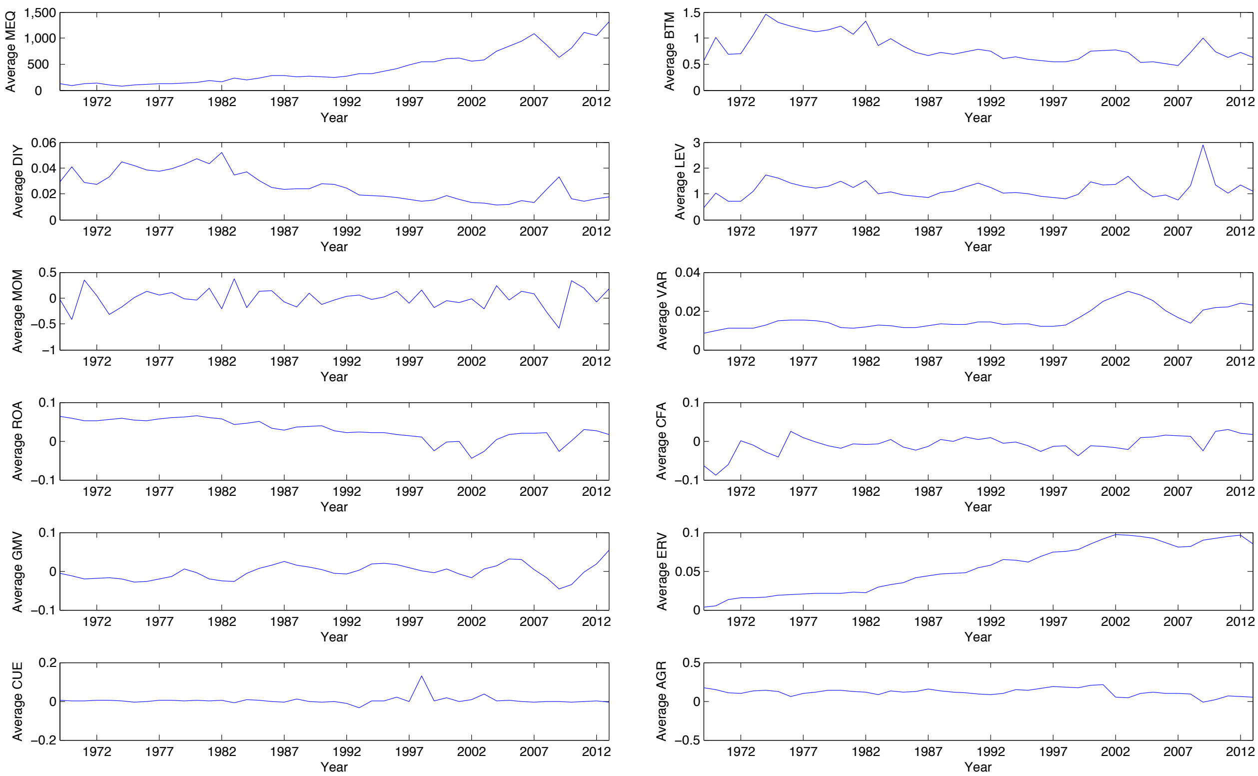

Figure 1: **Time-series of averages of firm characteristics.** This figure shows the cross-sectional average values of all analyzed firm characteristics from June 1969 to June 2013. A list of all abbreviations can be found in Appendix B. The calculation is done on an annual basis at the end of June to ensure that the retained information of the corresponding annual reports is available. All reported average values are calculated across all sample companies every year. *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.

In Table 2 we compute 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. It can be seen 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).

Char	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR
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
AGR												1

Table 1: **Correlation of firm characteristics.** The table displays the correlation of the firm characteristics within the cross-section. A list of all abbreviations can be found in Appendix B. *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.

Lag	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR
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

Table 2: **Autocorrelation of firm characteristics.** The table displays the autocorrelation of the regarded firm-characteristics. Each year, we compute the Pearson correlation of one firm characteristic of all firms with the same characteristic of all firms lagged by one to five years. The correlation is calculated only for the firms with available data. The correlations are then averaged over all sample dates. A list of all abbreviations can be found in Appendix B. The autocorrelations are calculated for the following 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.

3.2 Single characteristic portfolios

3.2.1 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 will consider two cases. The first one is defined with $\delta_T = \infty$ implying that no constraint is used for the optimization (*unconstrained policy*). The second case is set with $\delta_T = N_T^{-1}$ allowing for an intermediate level between no constraint and maximum constraint (*constrained policy*). For each of these two cases, we will look at the impact of risk aversion on the performance of the portfolio. Accordingly, we will 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. As in Brandt et al. (2009), we use a calibration sample of $\tau = 10$ years. We will study the sensitivity of the results to variations in τ in Section 4.2. We truncate our sample before 1969 in order to obtain a sufficient number of companies for which many characteristics are available. This means that the allocation starts at the beginning of 1979 and ends in June 2013.

Our eight key indicators consist of performance, turnover and leverage based measures. Transaction costs are modelled 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 the one of Brandt et al. (2009), we assume that transaction costs in 1979 are five times larger than those in 2013 (i.e. $A_{1979} = 5$ and $A_{2013} = 1$). Hence, the cost of transaction 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

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

Further we calculate the turnover as

$$Turn = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_t} |w_{i,T} - w_{i,T-}|.$$

On the side of the performance indicators we introduce volatility (Vol) as the standard deviation of annual returns. The Sharpe ratio (SR) is equal to the annualized return minus the risk free rate[10], divided by the portfolio's annualized volatility. 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). Figures are presented in bold font when the corresponding p-value is smaller than 10%. The CAPM alpha (α) and beta (β) result from regressing the portfolio returns against a value-weighted index consisting of all available stocks in the universe.

Lastly, we provide the average (over all dates) sum (SNW) and average proportion of negative weights (PNW) of the portfolios:

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

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

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

The sum of negative weights is a proxy for leverage because it represents the proportion of shortsales within the portfolio. An investor subject to Regulation T requirements must have this indicator below 0.5, which, as we will see, is not always the case for unconstrained policies.

Table 3 provides the results of the introduced key indicators which are obtained for the equally-weighted policy.

	Vol	SR	Turn	TC (%)	α	β	SNW	PNW
Equally-weighted benchmark	0.17	0.39	0.29	0.55	0.07	1.07	0.00	0.00

Table 3: **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 B.

3.2.2 Unconstrained policies

Table 4 shows the results of the unconstrained single characteristic portfolio policy. We report the eight key indicators presented in the previous subsection. Please note that this format will be the same for all subsequent tables in the current section. We further provide a list of all abbreviations in Appendix B.

The first and probably most obvious feature is that our results depend strongly on risk aversion. As expected, volatility decreases with γ: it is divided by a factor of 1.4 to 3.5 (across the 12 firm characteristics) when switching from $\gamma = 1$ to $\gamma = 10$. Turnover and transaction costs, too, are much impacted by risk aversion (divided by a factor of 4 to 10), as well as the proportion of negative weights (reduced by 50% to 95%). The sum of negative weights indicates that low risk aversion leads to very leveraged portfolios: the portfolios require from 50% to 380% of short sales. Nevertheless, we do not report a clear monotonous impact on the Sharpe ratio. The loss in Sharpe ratio for portfolios built on market equity (MEQ) is severe, from 1.15 for $\gamma = 1$ to 0.10 for $\gamma = 10$, while for portfolios designed upon dividend yields, it is the opposite (0.24 versus 0.52). For some characteristics (VAR, CFA, GMV, ERV), the maximum out-of-sample Sharpe ratio is attained for $\gamma = 5$.

When comparing the Sharpe ratios in Table 4 with those of the equally-weighted benchmark we find that only four values, namely for MEQ, BTM, DIY and AGR, are significantly higher than those of the benchmark.[11] Four characteristics seem little, especially given the high Sharpe ratios associated to low risk aversion. The lack of significance can only stem from the very high volatilities of the corresponding portfolios. Overall, a low risk aversion leads to unrealistic portfolios: for instance, one of them (CFA) has a negative beta (-0.64) and notably high turnover (10.74). In fact, we find a relationship between the characteristics which generate

[11] 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.

	Vol			SR			Turn			TC (%)		
γ	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
MEQ	0.52	0.19	0.17	**1.15**	0.42	0.10	4.18	0.59	0.46	7.54	1.01	0.82
BTM	0.56	0.21	0.18	**0.88**	0.46	0.24	4.71	0.63	0.49	9.12	1.22	0.88
DIY	0.22	0.16	0.16	0.24	0.50	**0.52**	2.85	0.77	0.75	5.32	1.53	1.49
LEV	0.62	0.20	0.17	0.65	0.42	0.25	3.58	0.54	0.52	7.04	0.98	0.93
MOM	0.69	0.22	0.18	0.71	0.37	0.15	9.28	1.02	1.13	17.59	1.97	2.27
VAR	0.49	0.15	0.14	0.35	0.49	0.44	2.88	0.76	0.70	4.96	1.45	1.38
ROA	0.59	0.19	0.18	0.64	0.40	0.23	4.81	0.76	0.81	9.77	1.52	1.62
CFA	0.47	0.16	0.16	0.35	0.44	0.38	10.74	2.69	2.55	23.88	6.36	5.82
GMV	0.33	0.18	0.19	0.13	0.30	0.29	7.52	1.45	1.55	12.57	3.13	3.13
ERV	0.35	0.15	0.15	0.31	0.42	0.37	2.22	0.68	0.74	3.97	1.24	1.36
CUE	0.26	0.17	0.17	0.32	0.39	0.39	3.96	1.10	0.90	8.42	2.31	1.79
AGR	0.55	0.20	0.18	**0.99**	0.50	0.23	9.71	0.93	0.94	19.54	1.89	1.83

	α			β			SNW			PNW		
γ	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
MEQ	0.56	0.08	0.02	1.79	1.07	0.98	-3.32	-0.08	-0.08	0.31	0.08	0.12
BTM	0.45	0.09	0.04	2.08	1.18	1.06	-2.11	-0.04	-0.03	0.50	0.08	0.03
DIY	0.07	0.09	0.10	0.69	0.85	0.87	-0.91	-0.14	-0.11	0.30	0.18	0.16
LEV	0.35	0.08	0.05	2.00	1.13	1.02	-1.76	-0.01	-0.07	0.60	0.02	0.04
MOM	0.44	0.08	0.03	2.05	1.17	1.06	-2.96	-0.10	-0.12	0.45	0.09	0.11
VAR	0.17	0.09	0.08	1.06	0.76	0.73	-1.16	-0.21	-0.21	0.41	0.15	0.11
ROA	0.36	0.08	0.04	1.53	1.09	1.03	-2.24	-0.08	-0.14	0.42	0.08	0.08
CFA	0.26	0.10	0.08	-0.64	0.58	0.73	-2.85	-0.74	-0.68	0.38	0.19	0.19
GMV	0.08	0.06	0.06	0.39	0.90	0.97	-2.48	-0.39	-0.44	0.36	0.17	0.21
ERV	0.12	0.08	0.07	0.94	0.80	0.78	-0.51	-0.10	-0.15	0.42	0.03	0.04
CUE	0.08	0.07	0.07	1.13	1.01	0.99	-0.91	-0.14	-0.09	0.13	0.04	0.03
AGR	0.53	0.10	0.04	1.31	1.10	1.07	-3.81	-0.12	-0.09	0.38	0.08	0.08

Table 4: **Unconstrained single characteristic portfolio policy.** This table displays the results for all key indicators and different risk aversions of the unconstrained single characteristic based portfolio policy. A list of all abbreviations can be found in Appendix B. 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 by using different risk-aversion inputs (γ) of 1, 5 and 10 for each firm characteristic. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally weighted benchmark.

high turnover and those which have the lowest autocorrelations in Table 2: MOM, CFA, GMV and AGR.

A second salient conclusion is that the performance of the regarded portfolios are driven by the choice of the underlying characteristic. If we look at the intermediate level of risk aversion ($\gamma = 5$), we see that volatility will range between 0.15 and 0.22, Sharpe ratio between 0.30

and 0.50 and turnover between 0.54 and 2.69. These discrepancies are strongly magnified if we consider a low risk aversion.

3.2.3 Constrained policies

So far our observations of the unconstrained case has shown that the choice of risk aversion strongly impacts the results of the portfolio policy. Low levels of risk aversion lead to relatively aggressive portfolio policies with high leverage, high volatility and large deviations from the market portfolio. On the other hand, a high risk aversion leads to lower levels of leverage, volatility and turnover. Given the purpose of our constraint we expect the overall level of leverage to decrease and to obtain smaller portfolio turnovers. Further, we do not expect Sharpe ratios to increase when applying the constraint. At first this might seem counter-intuitive since Jagannathan and Ma (2003) show that constraints can help to improve Sharpe ratios if optimization estimates are noisy. However, given that the approach by Brandt et al. (2009) is based on characteristics instead of returns this problem is lowered since firm characteristics are more stable over time. Consequently, the introduction of our constraint is not necessarily expected to lead to better out-of-sample Sharpe ratios.

In Table 5 we gather the results obtained when introducing the constraint of Equation 3 with $\delta_T = N_T^{-1}$. The first and probably most striking observation is that the overall results of the portfolio policies seem less volatile and more robust to changes in risk aversion compared to the unconstrained case.[12] The discrepancies in Sharpe ratio are much smaller and for half of the characteristics (DIY, VAR, CFA, GMV, ERV and CUE), the maximum difference is smaller than 0.1 in absolute value across all risk aversions. The largest spread in volatility is 0.04 in absolute value. These limited variations translate to CAPM alphas and betas. The turnover reaches reasonable values (never above 100%) even for low risk aversion (apart for the policies based on momentum). Consequently, apart for one exception (MOM and $\gamma = 1$), transaction costs are all below 20 basis points. Looking at portfolio weights, we observe that only one case (MEQ with $\gamma = 1$) displays more than 10% of negative weights, and the corresponding leverage is 13%, which is much lower than the 332% of the unconstrained case.

While the constraint reduces the spreads in performance for a given characteristic across different levels of risk aversion, it also curtails the discrepancies in the cross-section of characteristics. For $\gamma = 5$, the minimum Sharpe ratio is 0.34 and the maximum 0.45 and the spread in alpha is bounded by 0.02 in absolute value (0.07 in the unconstrained case). For low risk aversion, the reduction of spreads is even more apparent: from 1.02 to 0.23 for the Sharpe ratio and from 8.52 to 0.71 for turnover. This strong reduction in dispersion compared to the unconstrained policies holds for all eight indicators we report.

This reduced discrepancy across characteristics 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. The reduction is 50% on average, but it is usually larger

[12] Using a different optimization scheme, Grauer and Shen (2000) obtain similar results.

γ	Vol			SR			Turn			TC (%)		
	1	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
MEQ	0.19	0.19	0.16	**0.57**	0.40	0.18	0.63	0.50	0.31	1.11	0.86	0.52
BTM	0.22	0.20	0.18	**0.54**	0.43	0.26	0.69	0.55	0.45	1.28	1.07	0.79
DIY	0.16	0.16	0.16	0.40	0.42	0.44	0.59	0.52	0.56	1.06	0.98	1.04
LEV	0.21	0.19	0.17	0.47	0.40	0.27	0.56	0.48	0.47	1.01	0.88	0.82
MOM	0.20	0.19	0.17	0.50	0.36	0.23	1.18	0.68	0.79	2.17	1.30	1.46
VAR	0.18	0.15	0.15	0.40	**0.45**	0.42	0.56	0.51	0.48	1.05	0.92	0.88
ROA	0.18	0.18	0.17	0.46	0.39	0.29	0.62	0.59	0.51	1.14	1.12	0.96
CFA	0.18	0.18	0.18	0.40	0.37	0.35	0.72	0.65	0.72	1.28	1.25	1.35
GMV	0.17	0.17	0.17	0.34	0.34	0.34	0.60	0.57	0.60	1.07	1.03	1.08
ERV	0.16	0.15	0.15	**0.45**	**0.45**	0.41	0.47	0.47	0.50	0.79	0.81	0.86
CUE	0.17	0.17	0.16	0.37	0.37	0.37	0.51	0.49	0.49	0.90	0.88	0.89
AGR	0.19	0.18	0.17	**0.50**	0.44	0.27	0.90	0.62	0.69	1.59	1.15	1.26

γ	α			β			SNW			PNW		
	1	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
MEQ	0.11	0.08	0.03	1.10	1.07	0.99	-0.13	-0.04	0.00	0.14	0.05	0.00
BTM	0.11	0.08	0.05	1.26	1.17	1.05	-0.02	-0.01	-0.02	0.08	0.04	0.03
DIY	0.07	0.07	0.08	0.96	0.95	0.95	-0.02	-0.02	-0.02	0.07	0.06	0.07
LEV	0.09	0.08	0.05	1.19	1.12	1.03	0.00	-0.00	-0.05	0.00	0.00	0.04
MOM	0.10	0.07	0.04	1.11	1.10	1.03	-0.06	-0.03	-0.05	0.10	0.04	0.07
VAR	0.07	0.08	0.07	1.04	0.89	0.86	-0.02	-0.04	-0.05	0.03	0.05	0.05
ROA	0.08	0.07	0.05	1.06	1.06	1.02	-0.04	-0.04	-0.03	0.05	0.04	0.03
CFA	0.07	0.07	0.06	1.09	1.06	1.05	-0.04	-0.03	-0.05	0.06	0.04	0.06
GMV	0.06	0.06	0.06	1.00	1.01	1.03	-0.04	-0.03	-0.04	0.05	0.04	0.05
ERV	0.08	0.08	0.07	0.96	0.90	0.86	-0.00	-0.02	-0.04	0.00	0.01	0.02
CUE	0.07	0.07	0.07	1.03	1.01	1.00	-0.02	-0.02	-0.02	0.01	0.01	0.01
AGR	0.09	0.08	0.05	1.08	1.06	1.05	-0.08	-0.03	-0.03	0.08	0.04	0.04

Table 5: **Constrained single characteristic portfolio policy.** This table displays the results for all key indicators and different risk aversions of the constrained single characteristic based portfolio policy using constraint (3). A list of all abbreviations can be found in Appendix B. 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 by using different risk-aversion inputs (γ) of 1, 5 and 10 for each firm characteristic. 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.

for small values of γ. The average proportion of negative weights as well as total short-sales are also much smaller when introducing the constraint, which makes constrained policies easier to implement for financial institutions with leverage restrictions.

Surprisingly, the number of policies which significantly outperform the benchmark is higher after the introduction of the constraint (6) than before (4). This is a by-product of the reduction of the risk of all policies. In fact, this risk reduction generated by the constraint makes it possible for the investor to lower his risk aversion in the new setting. For most policies, this yields both higher alpha and lower leverage.

3.3 Double characteristic portfolios

Having analyzed 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 for the intermediate risk aversion ($\gamma = 5$).

3.3.1 Unconstrained policies

Table 6 shows the results for the unconstrained policies. The structure of the table follows the ones of Section 3.2. One first observation refers to the frequency of the most often occurring single firm characteristics within the double sorted portfolio policies. We see that VAR appears three times while MEQ, ERV, DIY and AGR appear twice in the top six pairs and MOM and GMV are represented three times in the bottom six combinations. This is in line with the figures of Table 4: DIY, VAR and AGR had the highest Sharpe ratio (for $\gamma = 5$), while MOM and GMV displayed the lowest ones.

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. This is further confirmed by the higher alphas. Turnover, transaction costs and the proportion of negative weights are most of the time (always when $\gamma = 5$ or $\gamma = 10$) lower for the top six policies.

The impact of the risk aversion parameter is substantial on all of the indicators we report. The magnitudes are close to those of Table 4. Moreover, as in the previous section, we acknowledge that the choice of characteristic is crucial for unconstrained portfolios. The policy based on the Fama French attributes (size and book-to-market) ranked 11th out of 66 for the Sharpe ratio for moderate risk aversion.

3.3.2 Constrained policies

As in the previous section, Table 7 completes the analysis with the constrained portfolio policies. As done before we show the results for the top and bottom six Sharpe ratio portfolios associated to a γ of five.

We first underline that four pairs (BTM-ERV, MEQ-DIY, MEQ-VAR and VAR-ERV) remain in the top six after the introduction of the constraint. The most represented characteristics

	γ	Vol 1	Vol 5	Vol 10	SR 1	SR 5	SR 10	Turn 1	Turn 5	Turn 10	TC (%) 1	TC (%) 5	TC (%) 10
top SR	BTM-ERV	0.62	0.16	0.14	**0.94**	**0.73**	**0.45**	4.97	1.08	0.88	9.52	2.14	1.71
	VAR-AGR	0.54	0.16	0.14	**0.98**	**0.71**	0.40	11.77	1.57	1.16	25.39	3.67	2.49
	MEQ-VAR	0.51	0.17	0.14	**1.36**	**0.70**	0.31	5.38	0.99	0.78	10.01	1.90	1.48
	DIY-AGR	0.58	0.18	0.17	**1.00**	**0.69**	0.42	10.29	1.29	1.32	20.48	2.59	2.60
	MEQ-DIY	0.48	0.17	0.15	**1.45**	**0.69**	0.31	5.17	1.01	0.78	9.71	1.95	1.48
	VAR-ERV	0.49	0.16	0.15	0.50	**0.65**	**0.57**	3.27	1.03	1.00	5.78	2.21	2.20
bottom SR	ERV-CUE	0.42	0.14	0.15	0.16	0.32	0.29	4.97	1.07	1.26	9.37	2.00	2.39
	LEV-MOM	0.73	0.21	0.19	0.63	0.32	0.10	6.11	1.71	1.68	11.58	3.49	3.44
	MOM-GMV	0.73	0.22	0.19	0.54	0.31	0.15	11.80	1.93	1.71	21.91	4.05	3.64
	MOM-ROA	0.69	0.22	0.19	0.64	0.31	0.11	10.18	1.90	1.55	20.55	3.91	3.13
	CFA-GMV	0.51	0.18	0.18	0.21	0.30	0.25	16.93	2.82	2.57	34.41	5.98	5.46
	GMV-CUE	0.38	0.19	0.20	-0.02	0.23	0.25	10.33	1.86	1.88	18.34	3.94	3.83

	γ	α 1	α 5	α 10	β 1	β 5	β 10	SNW 1	SNW 5	SNW 10	PNW 1	PNW 5	PNW 10
top SR	BTM-ERV	0.54	0.13	0.08	1.83	0.85	0.73	-2.55	-0.23	-0.21	0.52	0.13	0.09
	VAR-AGR	0.57	0.13	0.08	0.32	0.67	0.72	-4.66	-0.45	-0.29	0.41	0.17	0.16
	MEQ-VAR	0.69	0.13	0.06	1.32	0.80	0.74	-4.32	-0.36	-0.21	0.35	0.19	0.15
	DIY-AGR	0.57	0.14	0.08	1.22	0.91	0.87	-4.11	-0.28	-0.25	0.39	0.22	0.18
	MEQ-DIY	0.69	0.13	0.06	1.31	0.88	0.83	-3.75	-0.27	-0.18	0.34	0.23	0.19
	VAR-ERV	0.24	0.12	0.10	1.09	0.74	0.69	-1.64	-0.27	-0.27	0.54	0.17	0.13
bottom SR	ERV-CUE	0.09	0.06	0.06	0.64	0.78	0.80	-0.92	-0.16	-0.25	0.43	0.06	0.09
	LEV-MOM	0.43	0.07	0.02	1.66	1.13	1.06	-2.38	-0.30	-0.30	0.51	0.19	0.16
	MOM-GMV	0.37	0.07	0.04	1.46	1.03	0.98	-4.17	-0.47	-0.41	0.47	0.22	0.21
	MOM-ROA	0.41	0.06	0.02	1.69	1.17	1.10	-3.47	-0.38	-0.30	0.45	0.17	0.16
	CFA-GMV	0.16	0.06	0.05	-0.02	0.87	0.98	-4.63	-0.73	-0.64	0.42	0.25	0.24
	GMV-CUE	0.05	0.06	0.06	-0.04	0.83	0.94	-2.85	-0.46	-0.49	0.36	0.18	0.21

Table 6: **Unconstrained double characteristic portfolio policy.** This table displays the results for all key indicators and different risk aversions of the unconstrained double characteristic based portfolio policy. The table shows the top six and bottom six combinatorial portfolios out of the possible 66 possible pairs with respect to the Sharpe ratio and $\gamma = 5$. A list of all abbreviations can be found in Appendix B. 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 by using different risk-aversion inputs (γ) of 1, 5 and 10 for each firm characteristic. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally weighted benchmark.

remain the same for the best strategies (DIY, MEQ, ERV and VAR) and the worst ones (GMV, MOM, ROA). As such, the constraint does not strongly alter the overall ordering of the pairs.

Compared to the best constrained single characteristic policies, the best pairs provide an improvement in alpha (0.01 to 0.02 gain) and Sharpe ratio (the spreads between maxima is

	γ	Vol 1	Vol 5	Vol 10	SR 1	SR 5	SR 10	Turn 1	Turn 5	Turn 10	TC (%) 1	TC (%) 5	TC (%) 10
top SR	BTM-ERV	0.21	0.17	0.15	**0.63**	**0.54**	0.36	0.72	0.62	0.55	1.30	1.14	0.98
	DIY-ERV	0.16	0.16	0.16	**0.55**	**0.53**	0.47	0.57	0.60	0.63	1.03	1.09	1.14
	MEQ-DIY	0.18	0.17	0.16	**0.60**	**0.53**	0.31	0.66	0.62	0.51	1.19	1.12	0.92
	DIY-VAR	0.17	0.15	0.15	**0.51**	**0.52**	0.42	0.62	0.58	0.56	1.16	1.05	1.00
	MEQ-VAR	0.18	0.16	0.15	**0.58**	**0.51**	0.31	0.61	0.58	0.49	1.10	1.03	0.88
	VAR-ERV	0.15	0.15	0.14	0.44	**0.51**	**0.49**	0.53	0.51	0.52	0.96	0.91	0.94
bottom SR	MOM-ROA	0.19	0.20	0.18	**0.53**	0.37	0.22	1.01	0.71	0.73	1.86	1.36	1.33
	LEV-GMV	0.21	0.19	0.17	0.46	0.37	0.26	0.57	0.62	0.58	1.02	1.12	1.06
	ROA-GMV	0.19	0.18	0.17	0.43	0.36	0.31	0.63	0.66	0.63	1.14	1.21	1.15
	MOM-GMV	0.20	0.19	0.17	0.48	0.35	0.25	1.16	0.75	0.78	2.12	1.38	1.42
	GMV-CUE	0.17	0.17	0.17	0.36	0.34	0.34	0.64	0.62	0.64	1.12	1.10	1.13
	CFA-GMV	0.19	0.18	0.18	0.33	0.32	0.30	0.72	0.72	0.77	1.27	1.32	1.41

	γ	α 1	α 5	α 10	β 1	β 5	β 10	SNW 1	SNW 5	SNW 10	PNW 1	PNW 5	PNW 10
top SR	BTM-ERV	0.13	0.10	0.06	1.15	0.98	0.89	-0.02	-0.03	-0.06	0.09	0.05	0.04
	DIY-ERV	0.10	0.09	0.08	0.93	0.91	0.89	-0.01	-0.02	-0.04	0.04	0.05	0.06
	MEQ-DIY	0.11	0.09	0.06	1.06	0.97	0.94	-0.12	-0.05	-0.02	0.14	0.11	0.07
	DIY-VAR	0.09	0.09	0.07	1.01	0.88	0.89	-0.01	-0.04	-0.06	0.06	0.07	0.07
	MEQ-VAR	0.11	0.09	0.06	1.07	0.94	0.89	-0.11	-0.07	-0.05	0.14	0.09	0.07
	VAR-ERV	0.08	0.09	0.08	0.91	0.85	0.82	-0.02	-0.04	-0.05	0.02	0.05	0.06
bottom SR	MOM-ROA	0.10	0.07	0.04	1.10	1.11	1.05	-0.05	-0.04	-0.05	0.09	0.05	0.07
	LEV-GMV	0.09	0.07	0.05	1.18	1.07	1.02	0.00	-0.03	-0.06	0.01	0.04	0.05
	ROA-GMV	0.08	0.07	0.06	1.09	1.06	1.01	-0.05	-0.05	-0.05	0.06	0.06	0.06
	MOM-GMV	0.09	0.07	0.05	1.11	1.06	1.02	-0.06	-0.04	-0.05	0.10	0.06	0.08
	GMV-CUE	0.06	0.06	0.06	0.99	0.99	1.01	-0.04	-0.03	-0.04	0.05	0.04	0.05
	CFA-GMV	0.06	0.06	0.05	1.13	1.08	1.06	-0.05	-0.05	-0.06	0.06	0.06	0.07

Table 7: **Constrained double characteristic portfolio policy.** This table displays the results for all key indicators and different risk aversions of the constrained double characteristic based portfolio policy. The table shows the top six and bottom six combinatorial portfolios out of the possible 66 possible pairs with respect to the Sharpe ratio and $\gamma = 5$. A list of all abbreviations can be found in Appendix B. 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 by using different risk-aversion inputs (γ) of 1, 5 and 10 for each firm characteristic. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally weighted benchmark.

equal to 0.09, notably 0.54 and 0.45). The improvement in Sharpe ratio can also be seen by the fact that nearly all top SR portfolios have statistically significant higher Sharpe ratios than the equally-weighted benchmark. However we do not find any strong decline for volatility or beta. The lowest volatilities and betas are always those associated to the ERV or VAR indicators, which is also true for single characteristic policies. With respect to turnover or transaction costs, apart when comparing for the momentum-based policy, the spreads are also negligible. Equivalently to the single characteristic portfolios, we find that both, the sum of negative weights and the proportion of negative weights, have significantly reduced with the introduction of our constraint.

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. When comparing with the overlapping pairs between Tables 6 and 7, we infer that the reductions in magnitudes are close to those of single characteristic policies.

Overall, we find that the performance of characteristics-based policies depend 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 of characteristic) and hence curtails the risk of ill-advised decisions. Moreover, the constraint is also able to reduce the overall leverage in the double characteristic framework and diminishes the reported transaction costs.

4 Sensitivity analysis and robustness checks

4.1 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 approach and make use of the numerous combinations at hand to understand if feeding the optimization with additional characteristics will improve the policies *overall*. While the Sharpe ratio is the most natural performance indicator (overwhelmingly used by practitioners and academics), it can be improved by taking transaction costs into account so as to reflect risk-adjusted returns more realistically. Accordingly, we follow Garleanu and Pedersen (2013) and net the gross Sharpe ratio of transaction costs. We define the transaction cost-adjusted Sharpe ratio (ASR) as the raw Sharpe ratio minus the ratio of transaction costs (TC) to volatility (the transaction costs penalize the numerator of the Sharpe ratio): $ASR = SR - TC/\sigma$. In Figure 2, we plot the empirical cumulative distribution function (cdf) of this adjusted Sharpe ratio across all combinations of two (thin line) and three (thick line) firm characteristics. In the first case, there are 66 points, while in the second one, 220.

The very short distance (constrained case in grey) or intertwining (unconstrained case in black) between the curves do not make a clear case for an overall superiority of triple characteristic policies. We illustrate this finding with a simple example: in the constrained double characteristic portfolio policy of Table 7, the top two choices for $\gamma = 5$ are BTM-ERV and DIY-ERV with Sharpe ratios of 0.54 and 0.53, respectively. 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

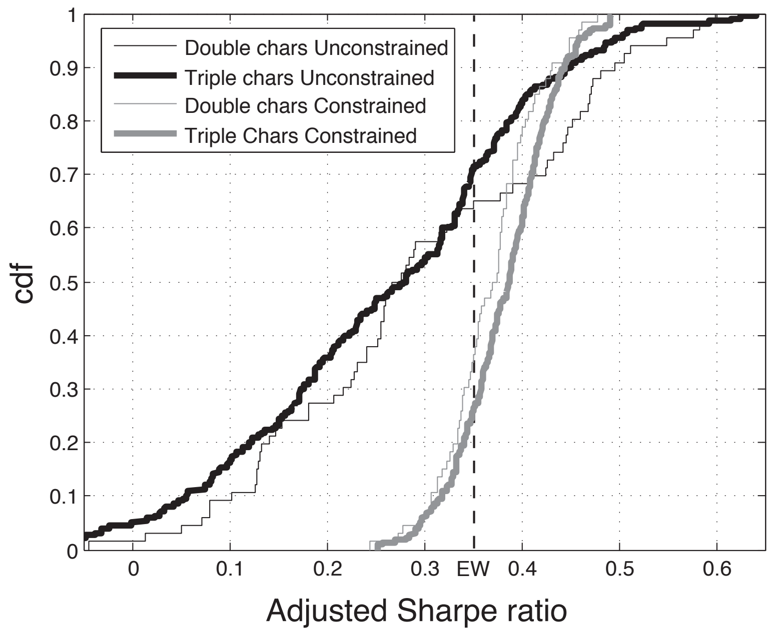

Figure 2: **Empirical distribution of the transaction cost-adjusted Sharpe ratio.** This figure shows the cumulative distribution function (cdf) of transaction cost-adjusted Sharpe ratio across all combinations of double and triple characteristic based portfolio policies. The unconstrained policies are plotted in black and the constrained ones in grey. We use thin lines for policies based on two characteristics and thick lines for polices based on three characteristics. The graph corresponds to $\gamma = 5$ and the calibration sample size is $\tau = 10$. *EW* represents the transaction cost-adjusted Sharpe ratio of the equally-weighted benchmark.

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, when increasing the number of characteristics per policy, 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 (7 times among the top 10 Sharpe ratios), MEQ (6 times) and DIY (5 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 further find that the triplet never outperforms the pair in terms of Sharpe ratio. Moreover the triplet always shows higher turnover and transaction costs. Accordingly, in our framework, an investor would be better off with only the size and book-to-market attributes. Consequently, it seems that compiling characteristics does not improve performance, apart for a few particular cases which can only be identified via data mining. Therefore, we believe that retaining 2 or 3 characteristics is a reasonable choice for investors.

Lastly, we recall that the transaction cost-adjusted Sharpe ratio of the equally-weighted benchmark is 0.35. We thus see in Figure 2 that a large majority (at least 75%) of the policies will outperform the benchmark when the optimization is constrained and $\gamma = 5$. However, this proportion falls to 30% without the constraint. Furthermore, we observe that the left tails of the distributions reach negative values for unconstrained policies. This is yet another illustration of the usefulness of the constraint: it allows to limit the risk of strong underperformance of the policies.

4.2 Sensitivity to sample size

In equation (4), we see that $\boldsymbol{\theta}_T$ 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 $\boldsymbol{\theta}_T$. This choice is somewhat arbitrary[13] and it is legitimate to wonder whether shorter sample sizes would lead to improvements or not. 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 to adjust $\boldsymbol{\theta}_T$ more rapidly, especially in times of turbulence.

Table 8 reports the impact on all eight key indicators when varying the estimation sample size, τ based on the double characteristic portfolio policy BTM-MEQ. As it is not tractable to provide the results for all combinations of all characteristics, we have performed the analysis on one double characteristic portfolio policy. 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 unconstrained portfolios (Panel A), we observe that the alpha is overall decreasing with τ. However, this relationship is not fully transposed to the Sharpe ratio. There is no monotonous impact on volatility or beta. However, turnover, transaction costs and leverage all decrease with τ. For turnover and transaction costs, this is quite straightforward, as longer sample size imply more stability of $\boldsymbol{\theta}_T$ through time.

When leverage constraints are enforced (Panel B), 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. The discrepancies in turnover, transaction costs and leverage are strongly reduced compared to the unconstrained policies. In short, this means that the constraint allows to maintain all indicators stable while increasing the Sharpe ratio when reducing the estimation sample. 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 $\boldsymbol{\theta}_T$ becomes progressively degenerate. For example, when $\tau = 2$, $\boldsymbol{\Sigma}_T$ in (6) is singular and it is possible to compute $\boldsymbol{\theta}_T$ using the Moore-Penrose inverse. This leads to a negative Sharpe ratio.

Overall, the unconstrained optimization can only generate relevant weights if the sample

[13] 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.

Panel A: Unconstrained Policies

τ	Vol	SR	Turn	TC (%)	α	β	SNW	PNW
	Key indicators							
10	0.17	**0.61**	1.20	2.40	0.11	0.93	-0.48	0.28
9	0.19	**0.62**	1.23	2.39	0.12	1.04	-0.48	0.28
8	0.19	**0.61**	1.31	2.57	0.12	1.02	-0.49	0.26
7	0.19	**0.60**	1.42	2.61	0.12	0.98	-0.51	0.25
6	0.20	**0.61**	1.52	2.69	0.13	1.00	-0.57	0.24
5	0.21	**0.65**	1.55	2.72	0.14	1.06	-0.63	0.24
4	0.19	**0.62**	2.03	3.59	0.13	0.87	-0.79	0.26
3	0.21	**0.70**	2.47	4.43	0.16	0.85	-0.95	0.29

Panel B: Constrained Policies

τ	Vol	SR	Turn	TC (%)	α	β	SNW	PNW
	Key indicators							
10	0.19	0.47	0.58	1.05	0.09	1.09	-0.05	0.09
9	0.19	**0.49**	0.59	1.06	0.09	1.11	-0.05	0.09
8	0.19	**0.50**	0.60	1.08	0.09	1.08	-0.04	0.08
7	0.19	**0.49**	0.62	1.12	0.09	1.07	-0.04	0.08
6	0.19	**0.49**	0.62	1.11	0.09	1.07	-0.04	0.08
5	0.20	**0.50**	0.63	1.14	0.10	1.13	-0.04	0.08
4	0.19	**0.52**	0.64	1.18	0.10	1.11	-0.04	0.07
3	0.19	**0.52**	0.67	1.22	0.10	1.10	-0.04	0.06

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 varying the estimation sample size (τ). 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 B. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally weighted benchmark.

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.

4.3 Adjusting the constraint to individual leverage restrictions

Table 9 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 all previous numerical applications so far, we have set the bindingness of the leverage constraint to $\delta_T = N_T^{-1}$ because it is a reasonable compromise between zero and the full constraint. In order to investigate the impact of variations in this parameter, we report the key indicators when $\delta_T = N_T^{-\kappa}$ and $\kappa \in (0.0, 0.1, \ldots 2)$ so that N_T ranges between N_T^{-2} and 1, which are

Panel A: $\gamma = 1$

κ	Vol	SR	Turnover	TC (%)	α	β	SNW	PNW
	Key indicators							
1.9	0.18	0.37	0.39	0.67	0.07	1.11	0.00	0.00
1.8	0.18	0.37	0.39	0.69	0.07	1.11	0.00	0.00
1.7	0.19	0.38	0.40	0.71	0.07	1.12	0.00	0.00
1.6	0.19	0.39	0.42	0.73	0.07	1.12	0.00	0.00
1.5	0.19	0.41	0.43	0.77	0.08	1.13	0.00	0.00
1.4	0.19	0.43	0.46	0.82	0.08	1.14	0.00	0.00
1.3	0.20	0.46	0.50	0.89	0.09	1.16	0.00	0.01
1.2	0.20	0.50	0.55	1.00	0.10	1.18	-0.01	0.03
1.1	0.21	**0.55**	0.63	1.14	0.11	1.20	-0.03	0.07
1.0	0.22	**0.61**	0.74	1.35	0.13	1.24	-0.07	0.13
0.9	0.23	**0.68**	0.89	1.62	0.15	1.28	-0.15	0.19
0.8	0.25	**0.76**	1.11	2.01	0.18	1.35	-0.29	0.24
0.7	0.27	**0.87**	1.40	2.55	0.22	1.41	-0.49	0.30
0.6	0.30	**0.96**	1.78	3.24	0.27	1.49	-0.77	0.33
0.5	0.35	**1.05**	2.29	4.20	0.34	1.56	-1.19	0.36
0.4	0.38	**1.11**	2.71	5.12	0.40	1.66	-1.49	0.38

Panel B: $\gamma = 5$

κ	Vol	SR	Turnover	TC (%)	α	β	SNW	PNW
	Key indicators							
1.9	0.18	0.35	0.37	0.65	0.06	1.11	0.00	0.00
1.8	0.18	0.36	0.38	0.66	0.07	1.11	0.00	0.00
1.7	0.18	0.36	0.38	0.67	0.07	1.11	0.00	0.00
1.6	0.18	0.37	0.39	0.69	0.07	1.11	0.00	0.00
1.5	0.19	0.37	0.40	0.71	0.07	1.11	0.00	0.00
1.4	0.19	0.39	0.42	0.74	0.07	1.11	0.00	0.00
1.3	0.19	0.40	0.44	0.79	0.07	1.11	0.00	0.01
1.2	0.19	0.42	0.48	0.85	0.08	1.10	-0.01	0.02
1.1	0.19	0.44	0.53	0.94	0.08	1.10	-0.02	0.05
1.0	0.19	0.47	0.58	1.05	0.09	1.09	-0.05	0.09
0.9	0.19	0.50	0.67	1.21	0.09	1.07	-0.09	0.15
0.8	0.19	**0.51**	0.75	1.38	0.10	1.06	-0.14	0.20
0.7	0.18	**0.52**	0.84	1.58	0.10	1.04	-0.21	0.22
0.6	0.19	**0.53**	0.88	1.68	0.10	1.04	-0.25	0.24

Table 9: **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 varying the intensity of the leverage constraint. We fixed $\delta_T = N_T^{-\kappa}$ for $\kappa \in (0.4, 0.5, \ldots, 1.9)$. The cases $\kappa = 0.4$ and $\kappa = 0.6$ are equivalent to the unconstrained policies for $\gamma = 1$ and $\gamma = 5$ respectively. A list of all abbreviations can be found in Appendix B. Bold figures indicate a statistically higher Sharpe ratio at a 10% confidence level, compared to the equally weighted benchmark.

the two bounds of Lemma 2.1.[14] As in the previous subsection, we proceed with the policies based on size and book-to-market. We provide the results for two levels of risk aversion ($\gamma = 1$ and $\gamma = 5$). The figures 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, and these stocks were therefore excluded from the optimization. In this case they were therefore excluded from the optimization scheme. As expected, the turnover, transaction costs, sum and proportion of negative weights decrease as the constraint becomes tighter (i.e. when κ increases). Further, the effect is more pronounced for $\gamma = 1$.

In the case of low risk aversion ($\gamma = 1$), both the Sharpe ratio and the volatility strongly decrease with κ. In contrast, in case of moderate risk aversion, only the Sharpe ratio displays a monotonous (decreasing) pattern (both the volatility and beta remain nearly constant). For both panels, subtracting transaction costs to the Sharpe ratio only marginally affects these conclusions. In Panel A, we find that for $\kappa = 1.3$, the policies are close to a long-only policy (less than 1% of shortsales) and the Sharpe ratio raises to 0.46, from 0.37 when $\kappa = 1.9$ (highly constrained program). Fine-tuning the constraint therefore allows to keep very low leverage while improving the risk-adjusted potential of the portfolio.

4.4 Factor tilts of different portfolio policies

When choosing a particular firm attribute over all others, the investor explicitly suggests that this attribute, or the underlying risk factor, is (at least to him) an important driver of returns in the cross-section. Consequently it seems reasonable that he expects to obtain a particular exposure to a factor related to this attribute. In this section, we focus our analysis on the size exposure (linked to market equity) and to the growth/value exposure (proxied by discrepancies in the book-to-market characteristic). The reason for this is that these two indicators are related to the ubiquitous Small-Minus-Big (SMB) and High-Minus-Low (HML) factors of Fama and French (1992) which are by now well known by investors. In order to quantify the exposures, we proceed as in Brandt et al. (2009) and compute the average exposures as

$$E_{SMB^*} = \frac{1}{T}\sum_{t=1}^{T} e_t^{MEQ}, \quad E_{BTM^*} = \frac{1}{T}\sum_{t=1}^{T} e_t^{BTM}, \tag{8}$$

where $e_t^{MEQ} = \boldsymbol{w}_t'\boldsymbol{x}_t^{MEQ}$ (*resp.* $e_t^{BTM} = \boldsymbol{w}_t'\boldsymbol{x}_t^{BTM}$) is the time-$t$ weighted average of the market equity (*resp.* book-to-market) firm characteristic. While the scores in (8) have little to do with the canonical loadings on SMB and HML, it is clear that they will highlight similar style exposures: a positive E_{SMB^*} will denote a portfolio tilted towards large cap stocks and a positive E_{HML^*} will correspond to a value portfolio. Accordingly, it is only a small abuse of language to consider that the expression of (8) corresponds to factor exposures. For this reason, we mark the resulting factors of (8) with a star to indicate this semantic interference. Table 10 shows

[14] Please note that we do not report results for $\kappa < 0.4$ in Panel A and $\kappa < 0.6$ in Panel B because the values do not change beyond these thresholds.

the exposures for all single characteristic portfolios and for the the double characteristic policy relying on BTM-MEQ.

Panel A: Unconstrained Policies

		Characteristic portfolio policy												
	γ	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR	BTM-MEQ
SMB*	1	-9.60	-1.22	0.19	-1.08	-0.84	-0.55	-1.45	-0.65	-0.42	-0.47	0.07	-1.82	-6.45
	5	-0.22	-0.08	0.08	-0.07	-0.03	0.06	-0.02	0.02	-0.08	0.04	-0.03	-0.10	0.41
	10	0.95	0.06	0.07	0.06	0.07	0.14	0.16	0.11	-0.04	0.11	-0.05	0.11	1.27
HML*	1	1.36	7.37	0.38	3.99	1.94	-0.38	1.82	-0.12	0.83	-0.57	0.05	1.65	5.01
	5	0.05	0.56	0.30	0.20	0.06	0.10	0.02	-0.25	0.16	0.00	-0.02	0.11	0.81
	10	-0.12	-0.29	0.29	-0.27	-0.17	0.16	-0.21	-0.27	0.08	0.07	-0.03	-0.08	0.28

Panel B: Constrained Policies

		Characteristic portfolio policy												
	γ	MEQ	BTM	DIY	LEV	MOM	VAR	ROA	CFA	GMV	ERV	CUE	AGR	BTM-MEQ
SMB*	1	-0.94	-0.17	0.04	-0.10	-0.08	-0.04	-0.13	-0.06	-0.03	-0.06	0.00	-0.13	-0.66
	5	-0.12	-0.06	0.04	-0.05	-0.01	0.02	-0.01	0.00	-0.02	0.00	0.00	-0.05	-0.11
	10	0.62	0.06	0.04	0.05	0.05	0.05	0.08	0.04	0.00	0.04	-0.01	0.07	0.49
HML*	1	0.13	0.94	0.10	0.40	0.19	-0.02	0.15	0.01	0.06	-0.08	0.00	0.11	0.77
	5	0.03	0.40	0.13	0.14	0.02	0.04	0.01	-0.03	0.04	-0.02	-0.01	0.06	0.26
	10	-0.08	-0.27	0.15	-0.23	-0.12	0.05	-0.09	-0.05	0.01	0.03	-0.01	-0.05	-0.13

Table 10: **Factor exposures of different characteristics-based policies.** This table displays the average factor tilting of 13 portfolio policies with respect to the proxies of the Fama and French (1992) risk factors HML and SMB for different risk aversions. The calibration period is $\tau = 10$ years. A list of all abbreviations can be found in Appendix B.

A first striking result is that the sign of the exposure may vary with the risk aversion when staying with the same firm characteristic. For instance, both for constrained and unconstrained policies, the exposure to SMB* is negative (with the exception of DIY and CUE based policies) when $\gamma = 1$. This can be explained by the fact that small cap stocks are known to outperform large cap stocks in the long run (Fama and French (1992)). However, small cap stocks are also more volatile. Consequently, when risk aversion is low and the focus is on delivering return, the exposure to SMB* should be negative (i.e. small cap stocks should have larger weights, everything else equal). But when risk aversion is high and the focus is on minimizing risk, then the exposure to SMB* should be positive. For $\gamma = 1$, we observe that the constrained and unconstrained policies relying on the MEQ attribute have the lowest SMB* exposure and that the policies relying on BTM have the highest HML* exposure. These rankings are however altered for higher levels of risk aversion. For $\gamma = 1$, we also find that the policies based on both MEQ and BTM have large amplitudes in their exposures to SMB* and HML*.

The average values of the Table 10 can be complemented by a dynamic analysis of the relative importance given to each characteristic. We consider the double characteristic policy based on BTM-MEQ and monitor the values of e_t^{MEQ} and e_t^{BTM} over time. Figure 3 shows the corresponding values when plotting e_t^{MEQ} and e_t^{BTM} over time. We consider four cases: with and without the constraint as well as low ($\gamma = 1$) and intermediate risk aversion ($\gamma = 5$).

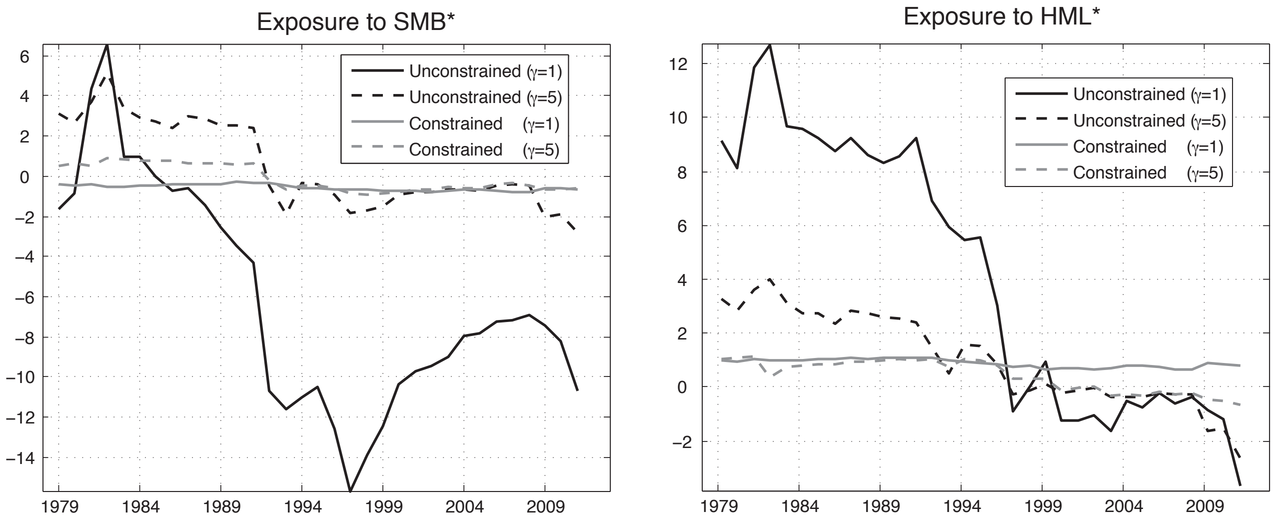

Figure 3: **Time-series of e_t^{MEQ} and e_t^{BTM} .** This figure plots the time-series of the values of $\boldsymbol{\theta}_T$ for the policies based on MEQ and BTM. On the left graph, we display the coefficient for the market capitalization (MEQ) characteristic and one the right graph, the coefficient for the book-to-market (BTM) characteristic. The unconstrained policies are plotted in black and the constrained ones in grey. We use thick lines for $\gamma = 1$, and dashed lines for $\gamma = 5$. The calibration sample size is $\tau = 10$.

When $\gamma = 1$ and the policy is unconstrained, we acknowledge a strong variation for both coefficients SMB* and HML*. When $\gamma = 5$, the variations follow the same patterns, but the trajectory is much smoother. The constrained case with $\gamma = 1$ is the most consistent of all since there is no variation in the sign of the score (negative for SMB* and positive for HML*). The negative exposures to HML* of the policies with intermediate risk aversion from 2000 onwards stand in contrast to Figure 12.1 in Ilmanen (2011). In this case the value premium in the US equity market is positive between 2000 and 2010. These negative exposures stem from the mitigated performance of value stocks compared to their growth counterparts in the 1990s and during the collapse of the internet bubble. We must therefore acknowledge that even though constrained policies are less subjects to variations in signs of risk premia associated to characteristics, they are not unconditionally immunized against the impacts of these shifts.

Overall, the smoothed patterns of constrained loadings are a gage of consistency: a rational investor does not expect his portfolio's exposures to the underlying firm characteristics (and, if they exist, to the corresponding risk premia) to evolve abruptly. This would only underline a lack of coherence of the policy. Accordingly, the constraint (or a higher risk aversion) is an efficient tool to ensure consistent risk exposures through time.

5 Conclusion

In this article we address several so far unsolved shortcomings of characteristics-based portfolio optimization. By introducing a novel leverage constraint into the modified framework of Brandt et al. (2009), we present an approach which aims at reducing the dispersion of weights around an agnostic prior: the equally-weighted portfolio. Based on an empirical analysis from 1969 to 2013 we find that portfolio policies including our constraint result in significantly lower short sales compared to the unconstrained policies. This makes constrained policies easier to implement, especially for investors with leverage constraints. Further, 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 seem less volatile and more robust. This stems from the fact that the discrepancies in Sharpe ratio are lowered when varying from high to low risk aversion. This is typical useful because the quantification of an investor's risk aversion is usually not straightforward. We further find that constrained policies have much lower levels of turnover and transaction costs compared to their unconstrained counterparts. Lastly, 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 characteristics which seem to generate value for the investor, 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 are likely to add any further value. Moreover, this latter attribute is unstable and often generates high 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. With respect to the number of possible characteristics which can be plugged into the optimization scheme, we find that beyond two characteristics, we cannot find any significant improvements.

Lastly, our approach is meant to remain flexible and the intensity of the constraint can be adjusted to fit to the investor's leverage target or to his risk budget. In a nutshell, this article enhances the applicability of characteristics-based portfolio choice and broadens the understanding of the underlying numerical optimization.

A 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. Based on this restriction, we calculate the firm characteristics at the end of June each year from 1964 to 2013 where the first years are solely used to calculate lagged returns. Hence, 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 return factors.

Following Brandt et al. (2009) we calculate the company's book equity as total assets minus total liabilities plus deferred taxes and investment tax credit minus the preferred stock value. Further, a company's market equity (MEQ) is determined as the price per share times the number of common shares outstanding. The book-to-market ratio BTM is defined as the book equity as defined above divided by the market equity. We calculate the current dividend yield DIY as total dividends divided by the number of common shares outstanding times the share price. The computation of leverage (LEV) follows Bhandari (1988): the leverage of a companyias 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 and total assets. The firm's cash-flow over assets (CFA) is calculated as net income plus depreciation minus the change in net working capital minus capital expenditures 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 minus costs of goods sold 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 prior 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. After the construction of all firm characteristics, we eliminate 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.

B Abbreviations

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	
α	Jensen's alpha based on a CAPM regression
β	Market beta based on a CAPM regression
PNW	Proportion of negative weights
SNW	Sum of negative weights
SR	Sharpe ratio
TC (%)	Transaction costs (in percent)
Turn	Portfolio turnover
Vol	Portfolio volatility

Input parameters	
γ	Risk aversion parameter
δ_T	Intensity of the leverage constraint
τ	Estimation sample size in years

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 A. 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 3.2.1. Finally, we specify input parameters as those variables which are varied for the understanding of the input sensitivities of the optimization algorithm.

C Proof of Proposition 2.1

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

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

and hence,

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

so that the first order condition implies that

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

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, $\boldsymbol{\theta}_T'\boldsymbol{x}_T'\boldsymbol{x}_T\boldsymbol{\theta}_T$ will 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 $\boldsymbol{x}_T'\boldsymbol{x}_T$, as in Appendix A in Coqueret (2014)).

D Proof of Lemma 2.1

We consider two opposite extreme configurations for the distribution of the y_{iT}. In the first scenario, there is one negative weight, $y_{1T} = -\sqrt{(N_T-1)\delta_T/N_T}$ and $N_T - 1$ positive weights $y_{jT} = \sqrt{\delta_T/(N_T(N_T-1))}$ for $j = 2, \ldots, N_T$. In this case, the minimum is equal to y_{1T}. This sequence satisfies the two equalities

$$\sum_{i=1}^{N_T} y_{iT} = 0 \quad \text{and} \quad \sum_{i=1}^{N_T} y_{iT}^2 = \delta_T.$$

Under these two constraints, we show below that it is impossible to find a minimum which is smaller than $-\sqrt{(N_T-1)\delta_T/N_T}$. We note S_+ for the subset of indices j such that $y_{jT} > 0$ and y_{T-} for the minimum value of $\boldsymbol{y}_T$. It must therefore hold that

$$\sum_{j\in S_+} y_{jT} \geq -y_{T-} \quad \text{and} \quad y_{T-}^2 \leq \delta_T - \sum_{j\in S_+} y_{jT}^2. \tag{9}$$

Since we want y_{T-} to be as large as possible, we want to allocate the positive weights in the most efficient manner (so that their L^2-norm is minimal). The solution of the classical linearly constrained quadratic program

$$\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}\boldsymbol{b} = \left\{ \underset{\boldsymbol{x}}{\operatorname{argmin}}\, \boldsymbol{x}'\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{A}\boldsymbol{A}' \succ 0 \right\}$$

implies that under the first inequality in (9), $\sum_{j\in S_+} y_{jT}^2 \geq y_{T-}^2/\text{card}(S_+)$. This is obtained by taking $\boldsymbol{A} = [1\ldots1]$ (row vector with length $\text{card}(S_+)$) and $\boldsymbol{b} = -y_{T-}$. From the second inequality in (9), we infer

$$y_{T-}^2 \leq \frac{\delta_T}{1 + 1/\text{card}(S_+)}, \tag{10}$$

which is maximal for $\text{card}(S_+) = N_T - 1$ (i.e., apart for its minimum value, $\boldsymbol{y}_T$ has only positive values). Moreover, in (10), the equality is reached when all positive weights have the same value. Accordingly, the minimal value for $y_{T-} = -\sqrt{(N_T-1)\delta_T/N_T}$.

The second scenario is the opposite situation where $y_{1T} = \sqrt{(N_T-1)\delta_T/N_T}$ and $y_{jT} = -\sqrt{\delta_T/(N_T(N_T-1))}$ for $j = 2, \ldots, N_T$. In this case, the minima are the y_{jT} for $j \geq 2$ and they can be proven to be maximal using the same technique as in the first case.

References

Arnott, R. D., J. Hsu, and P. Moore (2005). Fundamental indexation. *Financial Analysts Journal 61*(2), 83–99.

Asness, C., A. Frazzini, and L. H. Pedersen (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., M. Wauters, and D. Ardia (2014). Characteristic-based equity portfolios: Economic value and dynamic style-allocation. Working paper available at SSRN.

Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies 22*(9), 3411–3447.

Campbell, J., A. Lo, and A. MacKinlay (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.

Coqueret, G. (2014). Diversified minimum variance portfolios. *Annals of Finance*, forthcoming.

DeMiguel, V., L. Garlappi, F. J. Nogales, and R. Uppal (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science 55*(5), 798–812.

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

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

Fan, J., J. Zhang, and K. Yu (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association 107*(498), 592–606.

Garleanu, N. and L. H. Pedersen (2011). Margin-based asset pricing and deviations from the law of one price. *Review of Financial Studies 26*(6), 1980–2022.

Garleanu, N. and L. H. Pedersen (2013). Dynamic trading with predictable returns and transaction costs. *Journal of Finance 68*, 2309–2340.

Grauer, R. R. and F. C. Shen (2000). Do constraints improve portfolio performance? *Journal of Banking & Finance 24*(8), 1253–1274.

Grossman, S. J. and J.-L. Vila (1992). Optimal dynamic trading with leverage constraints. *Journal of Financial and Quantitative Analysis 27*(02), 151–168.

Hand, J. R. and J. Green (2011). The importance of accounting information in portfolio optimization. *Journal of Accounting, Auditing & Finance 26*(1), 1–34.

Hjalmarsson, E. and P. Manchev (2012). Characteristic-based mean-variance portfolio choice. *Journal of Banking & Finance 36*(5), 1392–1401.

Ilmanen, A. (2011). *Expected returns: an investor's guide to harvesting market rewards*. John Wiley & Sons.

Jacobs, B. and K. Levy (2013). Leverage aversion, efficient frontiers, and the efficient region. *Journal of Portfolio Management 39*(3), 54–64.

Jacobs, B. and K. Levy (2014). Traditional optimization is not optimal for leverage-averse investors. *Journal of Portfolio Management 40*(2), 30–40.

Jagannathan, R. and T. Ma (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance 58*(4), 1651–1684.

Ledoit, O. and M. Wolf (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance 15*(5), 850–859.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*(1), 77–91.

Pflug, G. C., A. Pichler, and D. Wozabal (2012). The 1/ n investment strategy is optimal under high model ambiguity. *Journal of Banking & Finance 36*(2), 410–417.

Plyakha, Y., R. Uppal, and G. Vilkov (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. and S. Lobe (2010). Fundamental indexing around the world. *Review of Financial Economics 19*(3), 117–127.