Financial Markets and Portfolio Management (2019) 33:243–275
https://doi.org/10.1007/s11408-019-00333-4

Alpha forecasting in factor investing: discriminating between the informational content of firm characteristics

Lars Heinrich[1] · **Martin Zurek**[2]

Published online: 25 July 2019

Abstract

This paper applies a linear alpha forecasting framework to enhance commonly used factor investing strategies by taking into account the informational content and interaction effects of selected firm characteristics. To demonstrate conditions under which it is beneficial to deviate from equally weighted characteristics, we evaluate a comprehensive number of factor portfolios. We consider four single-factor portfolios with 14 different firm characteristics in total and a multifactor portfolio where all factors are included. Empirically, the strategies are analyzed with the S&P 500, the Stoxx Europe 600 and the Nikkei 225 index. In addition, we also examine the strategies' performance in a simulation experiment and investigate the properties of the information coefficient estimates as a measure of the informational content. The empirical results are consistent with the simulation results, which reveal that the overall portfolio performance can be improved in well-defined factor models with a high dispersion among the mean information coefficients of the firm characteristics. In contrast, the naïve combination shows a comparable or better performance in factor models with a small dispersion in informational content between firm characteristics.

Keywords Factor investing · Multifactor · Alpha forecasting · Stock screening · Z-score · Information coefficient · Optimal orthogonal portfolio

JEL Classification G11 · G12 · G15

✉ Martin Zurek
zurek@europa-uni.de

Lars Heinrich
lars.heinrich@wwasset.de

1 W&W Asset Management GmbH, Tambourstraße, 70806 Kornwestheim, Germany

2 Viadrina European University, Große Scharrnstraße 59, 15230 Frankfurt (Oder), Germany

1 Introduction

Passive investment strategies that are based on firm characteristics, such as price-earnings or dividend-yield ratios, have gained popularity under the names of smart beta or factor investing. These are rule-based portfolio construction methods with the objective of offering exposure to specific risk factors and which claim to outperform an underlying capitalization weighted benchmark portfolio. The 314 published factors documented by Harvey et al. (2016) and the postpublication decay findings of McLean and Pontiff (2016) provide reasonable explanations for why many investors try to diversify their factor investing risk by creating a portfolio consisting of several firm characteristics. For example, in contrast to the Fama-French proxy portfolio for value, which relies on only the book-to-market ratio, the MSCI value index uses three additional firm characteristics. Instead of breaking new ground by finding a novel factor model, we concentrate on the question of the appropriate methodology to combine multiple types of firm characteristics for a variety of given factor models.

Within the rapidly growing zoo of factors, a proper combination of different types of firm characteristics is especially important for multifactor portfolios. However, this might also be relevant for single-factor portfolios that rely on a combination of multiple firm characteristics. To rank stocks according to multiple characteristics simultaneously, it is common practice to combine them into a weighted average over the scores of the single characteristics. This method is known as the bottom-up approach. In the top-down approach, the single-factor portfolios are first constructed and then combined within the multifactor portfolio. Bender and Wang (2016) evaluate these two approaches and demonstrate that the bottom-up method benefits from taking into account the interaction effects between factors, while the top-down approach does not. Therefore, this article focuses on a bottom-up approach based on z-score screening.

A z-score calculates the distance between an observation and the average of all observations and quantifies this distance in units of standard deviations. Pachamanova and Fabozzi (2014) present z-score screening as one of the recent trends in equity portfolio construction analytics. In addition to its user-friendliness, z-score screening can be motivated by a linear alpha forecasting model, also known as the forecasting rule of thumb in accordance with Grinold (1994). By motivating factor investing as an investment in the optimal orthogonal portfolio described by MacKinlay and Pástor (2000), we will determine portfolio weights proportional to these alpha forecasts.

Grinold and Kahn (2000, p. 263) define forecasted alphas as the difference between conditional and unconditional expected returns which is closely related to a cross-sectional Fama-MacBeth (FM) regression on lagged firm characteristics. Lewellen (2015) shows that FM-based return forecasts correspond well with true expected returns and provide an effective way to combine multiple firm characteristics into a composite estimate of expected returns. While Lewellen (2015) explores the predictive power of the overall factor models, it remains an open question whether and under which conditions FM slopes can help to improve the performance of factor investing strategies. This paper aims to fill this gap by conducting a horse race between the Grinold and Kahn (GK) approach and a naïve equally weighted combination of firm characteristics.

Breaking down FM regression slopes into informational content, interactions, and z-scores allows the assessment of the difference in the informational value of firm characteristics. Our main goal is to gain insights into the feasibility of applying these differences within the combination of multiple firm characteristics in factor investing strategies. As the naïve combination depends only on z-scores, a comparison between both approaches reveals the potential improvement of taking into account informational content and interactions. Informational content is measured by the information coefficient[1] (IC), which is the correlation between the applied information and the future return, whereas interactions are determined by the correlations between the firm characteristics. From the active portfolio management perspective, Grinold and Kahn (2000) refer to the IC as a forecasting skill, although we prefer the term informational content because in a factor investing strategy, the chosen characteristics should not be dependent on the skill level of the investor.

In comprehensive simulation and empirical backtest studies, portfolios obtained from GK alpha forecasts are compared to portfolios that use the naïve combination approach. By applying the strategies to a selection of different factors and data samples, we demonstrate conditions under which it is feasible to improve the out of sample portfolio performance. We consider a growth, value, quality, momentum and a multifactor portfolio with all 14 firm characteristics of the single factors. The firm characteristics in the factor portfolios are defined according to an industry standard of MSCI factor portfolios. Including stocks from the S&P 500, Stoxx Europe 600, and the Nikkei 225 index, the results are evaluated for three different data samples.

Our simulation and empirical results acknowledge the role of the IC parameter as an important performance driver and show that the IC structure of the applied factor models determines the potential performance gains from the GK approach. Due to high estimation errors, in factors with a small number of significant characteristics, the GK implementation reveals itself to be a challenging task. However, in well-specified factor models with a high number of characteristics with significant mean ICs and a high dispersion between these parameters, the naïve combination approach can be outperformed by incorporating the differences.

The remainder of the paper is organized as follows. Section 2 introduces the linear alpha forecasting model and the simplified version of the naïve combination approach. Section 3 presents the data set and the applied firm characteristics. Section 4 provides the simulation experiment results, and we examine the properties of the IC parameter estimates in a controlled environment. Section 5 provides the results of the empirical backtest. This includes the results of the long-only, quantile and perfect foresight portfolios. Section 6 contains some concluding remarks.

[1] Grinold (1989) stresses the important role of the IC in the fundamental law of active management (FLAM), where he points out that a portfolio's information ratio equals IC times the square root of the opportunity set size. Complementary work on the FLAM can be found in Clarke et al. (2002), Buckle (2004), Ye (2008) and Ding and Martin (2018).

2 Alpha forecasting

The investment universe is divided into an active portfolio $\mathbb{A}$ with $i = 1, \ldots, N$ securities and a passive benchmark portfolio $\mathbb{B}$, which acts as a proxy for the unobservable market portfolio. Throughout the article, the term "return" on a given security denotes its return in excess of the riskless rate.[2] In accordance with Treynor and Black (1973), we split the portfolio weighting approach in a weighting decision about the active portfolio $a_{\mathbb{A}}$ and a decision about the combination of $a_{\mathbb{A}}$ with the benchmark portfolio $a_{\mathbb{B}}$. Since our main purpose is to compare the GK and the naïve weighting approach we will concentrate on the active weighting decision. We assume that the returns of $\mathbb{A}$ and $\mathbb{B}$ are uncorrelated and therefore $a_{\mathbb{A}}$ represents the optimal orthogonal portfolio, which can be combined with $a_{\mathbb{B}}$ to form the optimal tangency portfolio. MacKinlay and Pástor (2000) show that in a factor model with a missing risk factor and a restricted form for the residual covariance matrix[3] $\Phi = \sigma^2 I$, the weights of the optimal orthogonal portfolio are proportional to the securities' alpha parameters:

$$a_{\mathbb{A}} = \frac{\alpha}{\mathbf{1}'\alpha}. \tag{1}$$

This result provides a motivation to apply factor portfolios with characteristic-based weighting schemes. In accordance with Fama-French, we separate securities into a long and short portfolio. Stocks with positive alpha forecasts are assigned to the long portfolio and stocks with negative alpha forecasts are assigned to the short portfolio:

$$a_{\text{long}} = \frac{\hat{\alpha}^+}{\mathbf{1}'\hat{\alpha}^+} \quad \text{with} \quad \hat{\alpha}^+ = \max(\hat{\alpha}, 0), \tag{2}$$

$$a_{\text{short}} = \frac{\hat{\alpha}^-}{\mathbf{1}'\hat{\alpha}^-} \quad \text{with} \quad \hat{\alpha}^- = \max(-\hat{\alpha}, 0). \tag{3}$$

In addition to long-short (LS), we build long-only (LO) portfolios, wherein only the long portfolio is taken into account.

We assume that security returns in the active portfolio follow a multivariate conditional normal distribution. As conditioning information, an amount of M different types of stock characteristics, hereinafter referred to as signals, can be observed for each of the N companies. For simplification purposes, we assume that these signals only have an informational value for their specific company, which means that signals from different companies do not correlate with each other or with returns from other companies. Additionally, we assume that signals are uncorrelated with benchmark returns so that we can disregard benchmark timing. We follow the approach of Grinold and Kahn (2000, p. 263) and define the forecasted alpha of a security as the difference between the conditional and unconditional expected returns:

[2] Excess returns are marked with an asterisk.

[3] Note that Φ is diagonal and proportional to the identity matrix. Complementary information about this assumption can be found in MacKinlay and Pástor (2000, p. 887).

$$\alpha_{GK,i,t} = \mathrm{E}\left[R^*_{i,t}|g_{i,t-1}\right] - \mathrm{E}[R^*_{i,t}] = \Lambda'_{12}\,\Lambda^{-1}_{11}\left(g_{i,t-1} - \mathrm{E}\left[g_{i,t-1}\right]\right), \tag{4}$$

with

$$\begin{aligned}\mathrm{Corr}\left[g_{i,t-1}, g_{j,t-1}\right] &= 0\ \forall\ i, j = 1\ldots N \Leftrightarrow i \neq j,\\ \mathrm{Corr}\left[g_{i,t-1}, R_{j,t}\right] &= 0\ \forall\ i, j = 1\ldots N \Leftrightarrow i \neq j,\\ \mathrm{Corr}\left[g_{i,t-1}, R_{\mathbb{B},t}\right] &= 0\ \forall\ i = 1\ldots N.\end{aligned}$$

$R^*_{i,t}$ denotes the return of a security i at time t. $g_{i,t-1}$ is the $(M \times 1)$-vector of signal observations. The $(M \times M)$-matrix Λ_{11} stands for the cross-sectional covariance matrix of signal observations among all securities. Furthermore, Λ_{12} is a $(M \times 1)$-vector of covariances between the cross-sectional returns and lagged signal observations among all securities for each signal in $g_{i,t-1}$. For better interpretation, Eq. (4) can be further rewritten as:

$$\alpha_{GK,i,t} = \sigma_{R^*_t}\,k\,C^{-1}\,z_{i,t-1}, \tag{5}$$

with

$$k = \left(\mathrm{Corr}\left[R^*_t, g_{1,t-1}\right] \cdots \mathrm{Corr}\left[R^*_t, g_{M,t-1}\right]\right),$$

$$z_{i,t-1} = \begin{pmatrix} \frac{(g_{i,1,t-1} - \mathrm{E}[g_{1,t-1}])}{\sigma_{g_{1,t-1}}} \\ \vdots \\ \frac{(g_{i,M,t-1} - \mathrm{E}[g_{M,t-1}])}{\sigma_{g_{M,t-1}}} \end{pmatrix}.$$

In the cross-sectional setting, the parameter $\sigma_{R^*_t}$ represents a measure of return dispersion, but since it is a constant for all securities, it can be neglected. The correlations between lagged signals and returns in k are the ICs which measure the signals' informational content. $z_{i,t-1}$ is the vector of the standardized signals which are called z-scores. To avoid bias from industry-specific characteristics all characteristics are z-score standardized depending on their industrial specific cross-sectional expected values and standard deviations. For the securities' sector allocation, GICS sector classification codes are used. The inverse of the signals' correlation matrix C^{-1} is responsible for ensuring that highly correlated signals will have a lower impact on the alpha forecast and vice versa. Since we have chosen to disregard benchmark timing due to uncorrelated signals and benchmark returns, we can directly apply ICs to residual returns. With the given benchmark portfolios, we calculate residuals from the one-factor regressions of security returns against the benchmark returns. Because factor investing strategies are based on cross-sectional anomalies, parameter estimation is conducted with regard to cross-sectional observations. Therefore, k and C will be the same for all securities. This is equivalent to the assumption that a specific stock characteristic will have the same informational content and linear comovement with other signal types for all securities. To obtain ex-ante estimates, we will estimate ICs and C from the time series averages of their cross-sectional estimates.

In practice, smart beta products can differ in the way they translate scores into portfolio weights, but in most cases, ICs and correlations among signals are ignored. In fact, most methods simply use equally weighted z-scores:

$$\alpha_{NZ,i,t} = \sum_{m=1}^{M} \frac{1}{M} \frac{\left(g_{i,m,t-1} - \mathrm{E}\left[g_{m,t-1}\right]\right)}{\sigma_{g_{m,t-1}}}. \tag{6}$$

The same alpha forecasts for both methods can be obtained by disregarding scaling factors such as $\sigma_{R^*_{i,t}}$ or $\frac{1}{M}$ and by assuming identical ICs and uncorrelated signal types. Therefore, the naïve z-score (NZ) weighting method can be seen as a simplified method of alpha forecasting.

3 Data

With the S&P 500 (SPX), the Stoxx Europe 600 (STX) and the Nikkei 225 (NKY) index, we investigate the results for three different data samples. These indices currently account for 91.5%[4] of the MSCI World index and are commonly used as benchmark portfolios. As a proxy for the risk-free rate in the US market, we use the one-month T-bill rate. Because there are no comparable one-month treasury bills in the Japanese and European markets, we assume the three-month Japan Treasury Discount Bill and the three-month Euro Government Bond rate as risk-free rates for their corresponding markets.[5] In contrast to the SPX and NKY, the STX is affected by currency effects. To prevent the results for the STX sample from being affected by currency effects, we assume that currencies are fully hedged and measure stock returns in the local currency. Any hedging costs are not considered. Our data set[6] includes stock returns and observations of firm characteristics from the beginning of 2002 until the end of 2016. We decide against a longer time horizon due to data availability for the NKY and STX indices and because of observable change points within the return generating process in the US market starting in 2002/2003. Consistent with Green et al. (2017) we observe a higher number of significant predictors for monthly US returns prior to 2002. We partly address the question of model misspecification by also considering factor portfolios with a low number of significant predictors. However, due to limited space, we do not account for time-varying predictability.

We apply the well-known value, growth, momentum, and a quality factor which has become increasingly important over the course of the last few years. The selection of

[4] Based on the MSCI World weights reported by Bloomberg on the last backtest date (31.12.2016).

[5] Annualized monthly mean returns and standard deviations of the benchmark portfolios and risk-free rates are shown in Table 9 in the appendix.

[6] The data set has been provided by the Bloomberg database.

firm characteristics is based on MSCI factor portfolios,[7] which find wide acceptance in the industry. Since we are interested in the combination of firm characteristics, we choose factors with at least two signals. An overview of the factor strategies and corresponding firm characteristics is shown in Table 1.

Because some characteristics depend on stock prices and are therefore subject to daily fluctuations, we apply daily trailing data observations to ensure that our calculations are always based on the most recent data. For accounting numbers, the set of data files that were actually commercially available in the forecasting month, are used to calculate firm characteristics. Since stocks are separated between high and low factor exposures, the direction of the applied information signals has to be determined so that a higher z-score will imply higher returns and vice versa. In this sense, the observations for debt to equity and earnings variability are multiplied by -1. To ensure that forecasts are not distorted by outliers, firm characteristics are winsorized above the 95th percentile and below the 5th percentile.[8]

In Table 2, for each data sample, the time series averages of 180 cross-sectional IC estimates are reported. To quantify the uncertainties in the inferences, we calculate bootstrap standard errors (SE) and 95% confidence intervals over 100,000 resamples. The sample standard deviation of the average ICs, across bootstraps, is used as an estimate of the standard error. To obtain lower and upper limits of the confidence intervals for the mean ICs the 2.5% and 97.5% centiles of the bootstrap sample means are calculated. The mean parameters show that ICs vary around averages, which are very close to zero. Nevertheless, the majority of signals have a positive mean IC. Furthermore, characteristics in some factors exhibit highly different average ICs, which we will refer to as factors with high IC dispersion. For instance, EPSG1Y and EPSGT in the STX sample possess approximately three times the mean IC of EPSG3Y or IGR. Moreover, CFOtEV and EtP are both significant, with relatively high mean ICs of 2.9% and 2.0%, respectively, whereas the other two characteristics in the value factor have negative mean ICs. Therefore, the value and growth factor of the STX sample represent factor portfolios with a high dispersion in informational content between the different signals.
In contrast, the characteristics in the value factor for the NKY sample show a very low IC dispersion, because all characteristics are significantly positive and three of the four characteristics have very strong ICs of approximately 3.0% to 4.0%.

Table 2 shows that the applied multifactor models in our samples exhibit three different cases of predictability. With only three characteristics with significant mean ICs in the SPX sample, the SPX multifactor portfolio represents the situation where an investor has not identified the best ex-ante predictors in the factor zoo. In contrast, the multifactor model in the STX sample, which has nine significant characteristics, represents a well-defined factor model. In the NKY sample, the EPSG1Y, EVar, and all value characteristics are significant, which indicates a special situation where most

[7] Here we use the MSCI Global Investable Market Value and Growth Index Methodology (September 2017), the MSCI Quality Indexes Methodology (June 2017) and the MSCI Momentum Indexes Methodology (June 2017). Due to data availability, our factor specification differs from the MSCI standard in some respects, particularly in that we apply trailing instead of forward data.

[8] Descriptive statistics for the winsorized firm characteristics are shown in Table 10 with average firm characteristics separated for all sector groups in Table 11 in the appendix.

Table 1 Overview: factor strategies and firm characteristics

Factor	Characteristic	Definition
Growth	Earnings Growth 1 Year (EPSG1Y)	$\text{EPSG1Y}_t = \text{YoY EPS Growth}_t$
	Earnings Growth 3 Years Average (EPSG3Y)	$\text{EPSG3Y}_t = \frac{1}{3}\sum_{t=1}^{3} \text{YoY EPS Growth}_t$
	Internal Growth Rate (IGR)	$\text{IGR}_t = \frac{\text{TTM EPS}_t}{\text{BVPS}_t} - \frac{\text{TTM DPS}_t}{\text{BVPS}_t}$
	Earnings Growth Trend (EPSGT)	$\text{EPSGT}_t = \frac{\beta_{EPS}}{1/5\sum_{t=1}^{5} \lvert\text{EPS}_t\rvert}$
	Sales Growth Trend (SPSGT)	$\text{SPSGT}_t = \frac{\beta_{SPS}}{1/5\sum_{t=1}^{5} \lvert\text{SPS}_t\rvert}$
Value	Dividend Yield (DivYld)	$\text{DivYld}_t = \frac{\text{TTM DPS}_t}{\text{P}_t}$
	Earnings to Price (EtP)	$\text{EtP}_t = \frac{\text{TTM EPS}_t}{\text{P}_t}$
	Book to Price (BtP)	$\text{BtP}_t = \frac{\text{BVPS}_t}{\text{P}_t}$
	Cash Flow to Enterprise Value (CFOtEV)	$\text{CFOtEV}_t = \frac{\text{TTM CFO}_t}{\text{EV}_t}$
Quality	Return on Equity (RoE)	$\text{RoE}_t = \frac{\text{TTM EPS}_t}{\text{BVPS}_t}$
	Debt to Equity (DtE)	$\text{DtE}_t = \frac{\text{TD}_t}{\text{BV}_t}$
	Earnings Variability (EVar)	$\text{EVar}_t = \sqrt{1/5\sum_{t=1}^{5}\left(\text{YoY EPS-Growth}_t - \overline{\text{YoY EPS-Growth}_t}\right)^2}$
Momentum	6 Months Price Momentum (Pmom6M)	$\text{n-Months Price Momentum}_t = \frac{\left(\frac{P_{t-1}}{P_{t-n-1}}-1\right)-R_{f,t}}{\sigma_t}$
	12 Months Price Momentum (Pmom12M)	

With: *P* share price, *EV* enterprise value, *TD* total debt, *BV* book value, *BVPS* book value per share, *EPS* earnings per share, *SPS* sales per share , *DPS* dividend per share, *CFO* cash flow from operations, *YOY* year over year, Trailing 12 Months (TTM). σ_t is the annualized volatility of weekly returns over the last 3 years. For the calculation of EPSGT and SPSGT, the last 5 yearly restated EPS and SPS are used. β_{EPS} and β_{SPS} represent slope coefficients from regressions of yearly EPS and SPS observations against the number of months in the 5-year observation interval. For a calculation example of EPSGT and SPSGT, see the MSCI Global Investable Market Value and Growth Index Methodology (2017, p. 32)

of the informative predictors are concentrated in a single factor. With a mean IC of 4% in BtP, the NKY sample contains the characteristic with the highest value. Interestingly, CFOtEV is the only characteristic, that is significant in all three samples with almost the same mean IC of 3%. With the exception of the STX results, the

Table 2 Mean ICs and bootstrap results of the SPX, STX, and NKY samples

Factor	Characteristic	SPX				STX				NKY			
		Mean	SE	0.025-Qt.	0.975-Qt.	Mean	SE	0.025-Qt.	0.975-Qt.	Mean	SE	0.025-Qt.	0.975-Qt.
Growth	EPSG1Y	0.778	0.501	− 0.201	1.767	**2.710**	0.490	1.750	3.680	**2.960**	0.630	1.770	4.240
	EPSG3Y	0.402	0.434	− 0.452	1.251	**0.810**	0.410	0.100	1.620	− 0.050	0.670	− 1.090	1.520
	IGR	**1.321**	0.557	0.230	2.406	**1.140**	0.490	0.180	2.090	1.350	0.940	− 1.230	2.450
	EPSGT	1.243	0.699	− 0.122	2.620	**2.830**	0.740	1.380	4.290	0.700	0.830	− 0.570	2.680
	SPSGT	0.344	0.537	− 0.715	1.398	0.710	0.560	− 0.390	1.810	− 0.860	0.750	− 1.230	1.720
Value	DivYld	0.977	0.631	− 0.259	2.212	− 0.160	0.690	− 1.520	1.200	**2.350**	0.730	0.690	3.540
	EtP	0.576	0.661	− 0.723	1.872	**2.010**	0.600	0.830	3.180	**3.490**	0.820	1.460	4.650
	BtP	− 0.375	0.783	− 1.903	1.154	− 0.780	0.960	− 2.660	1.080	**3.950**	0.910	1.320	4.890
	CFOtEV	**2.896**	0.623	1.676	4.114	**2.870**	0.490	1.900	3.840	**2.980**	0.690	2.050	4.770
Quality	RoE	**1.528**	0.622	0.308	2.742	**2.620**	0.720	1.220	4.030	0.740	0.960	− 1.360	2.410
	DtE	0.600	0.443	− 0.271	1.461	**1.890**	0.580	0.760	3.020	0.700	0.890	− 0.690	2.790
	EVar	0.541	0.499	− 0.443	1.512	0.220	0.480	− 0.700	1.170	**1.760**	0.930	0.780	4.430
Momentum	Pmom6M	0.091	0.843	− 1.553	1.735	1.670	0.990	− 0.280	3.610	− 1.450	1.050	− 3.260	0.860
	Pmom12M	0.718	0.955	− 1.147	2.590	**3.520**	1.110	1.340	5.710	− 0.400	1.090	− 2.820	1.460

This table shows (in%) the mean ICs and bootstrap results of the 180 IC parameter estimates of the whole time interval of the SPX, STX and NKY samples. Standard errors and the 2.5 and 97.5 percentiles of the 95% confidence interval are drawn from 100.000 bootstrap resamples. Mean ICs where the confidence interval does not cross-zero are highlighted in bold momentum characteristics display no significant ICs and are even negative in the NKY sample.

Figures 1, 2 and 3 show heatmaps from time series averages of cross-correlations among the firm characteristics for each data sample. As a high number of correlation coefficients are smaller than 20%, most of the characteristics are not highly correlated. This indicates that taking multiple firm characteristics into account benefits from diversification effects. In some individual cases, signal pairs like RoE and IGR, as well as Pmom6M and Pmom12M, appear to have a high positive correlation. The higher correlation coefficients lead to a reducing effect on the informational content of the corresponding signals. But, since the majority of the correlation coefficients are small, the impact on the overall alpha forecasts is very limited. In Sect. 5.2, we investigate the maximum potential performance improvement by investigating the results in a perfect foresight scenario. The results highlight that taking into account the signal correlation matrix can improve performance, but the main source of performance contribution is found in the consideration of ICs. Therefore, in a comparison of the GK and NZ methods, the benefits of the NZ method should depend on how well the assumptions of identical ICs for the different signals are met. Similarly, the benefits of the GK method depend on the capability to estimate the ICs and the sensitivity of the portfolio weighting methodology to estimation errors. While investigating these propositions, we will compare the NZ and the GK method in the following simulation experiment and backtest study.

4 Simulation

Because under real conditions, the true return generating process (RGP) is unknown, model risks, such as missing risk factors, can influence the backtest results. The following out-of-sample simulation allows us to examine the precision and bias of IC estimation as well as its influence on portfolio performance in a controlled environment. To obtain a robust result, we evaluate 1,000 simulation samples. Because we simulate using an estimated model, the simulation results are equivalent to parametric bootstrap samples, to which we can apply standard errors and confidence intervals.

4.1 Model and parameter assumptions

We assume the RGP to follow a linear factor model conditioned on the 14 firm characteristics from Table 1. Since the true coefficients are known from our simulation assumptions, we can evaluate the actual portfolio performance. In order to get realistic parameter assumptions, we apply simulation parameters based on our return and standardized characteristic observations from the SPX, STX and NKY data samples described in Sect. 3. Returns for a security $i = 1, \ldots, N$ are generated from the following multifactor model:

$$\mathbb{R}^{*}_{i,t} = \alpha_{i,t} + \beta_i \, R^{*}_{\mathbb{B},t} + \varepsilon_{i,t}, \tag{7}$$

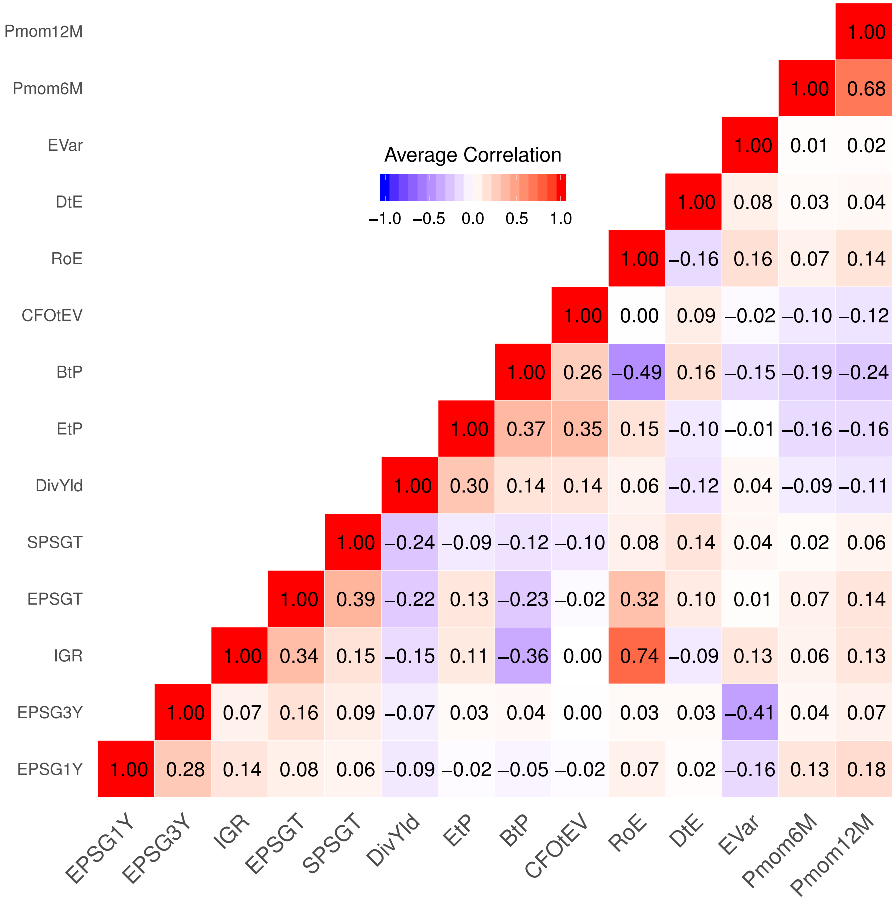

Fig. 1 Average correlations between firm characteristics in the SPX sample

with

$$\varepsilon_t \sim N(\mathbf{0}, \Psi),$$
$$R^*_{\mathbb{B},t} \sim N(\mu_{\mathbb{B}}, \sigma_{\mathbb{B}}).$$

Residuals $\varepsilon_{i,t}$ are simulated from a multivariate normal distribution with zero mean and covariance Ψ. Assumptions on Ψ are based on covariance estimations for residuals calculated from one-factor regressions of stock return observations against their according index returns. To overcome problems of large scale, shrinkage estimation is conducted by shrinking the empirical correlations toward the identity matrix with an analytic determination of the shrinkage intensity as described by Schäfer and Strimmer (2005).[9] Benchmark returns are assumed to be normally distributed with assumptions

[9] The simulation is conducted using the programming language R. Here, the cov.shrink function from the corpcor package is applied.

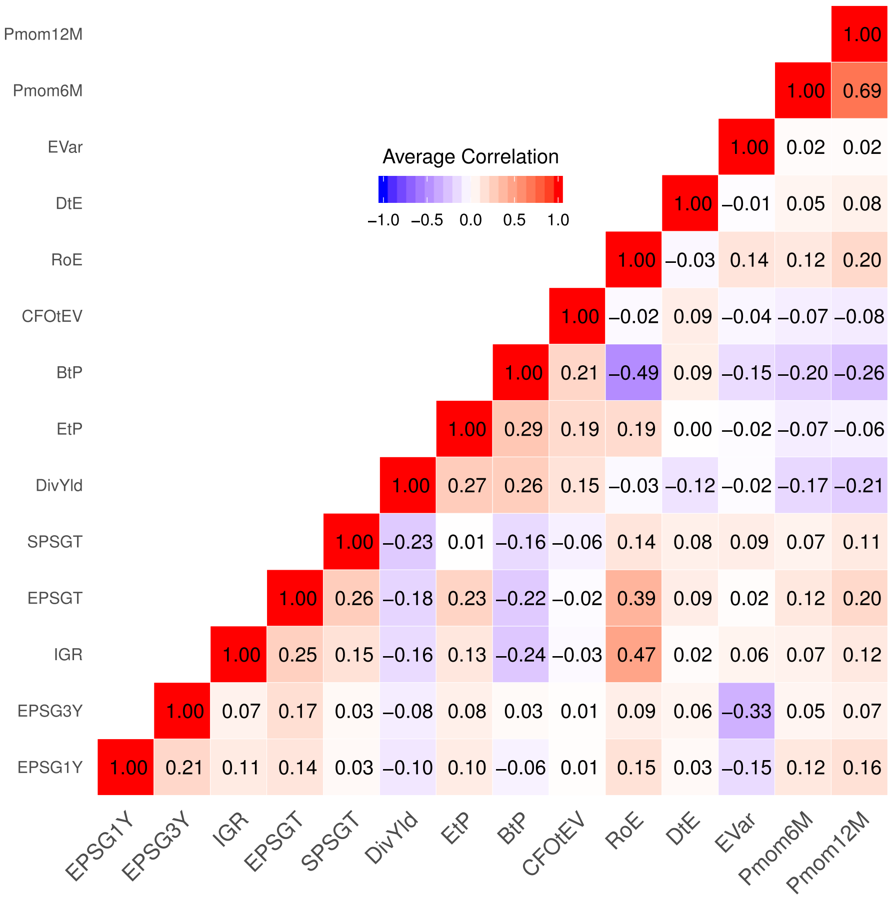

Fig. 2 Average correlations between firm characteristics in the STX sample

on $\mu_{\mathbb{B}}$ and $\sigma_{\mathbb{B}}$ based on maximum likelihood (ML) estimates of the index return observations. Parameters for β_i are evenly spread between 0.5 and 1.5, whereas $\alpha_{i,t}$ parameters are determined on conditioning information signals:

$$\alpha_{i,t} = \sigma_{\varepsilon_{i,t}} \, k_{cs,t} \, C_{cs,t}^{-1} \, z_{i,t-1}, \tag{8}$$

with

$$z_{i,t-1} \sim N\left(0, C_{cs,t}\right).$$

We assume the conditioning signals to follow a cross-sectional multivariate distribution. Due to the application of cross-sectional information, the signals' ICs in $k_{cs,t}$ and cross-sectional correlations $C_{cs,t}$ are the same for each security. In each of the 1,000 simulation samples, a subset of 180 parameter assumptions for $k_{cs,t}$ and $C_{cs,t}$

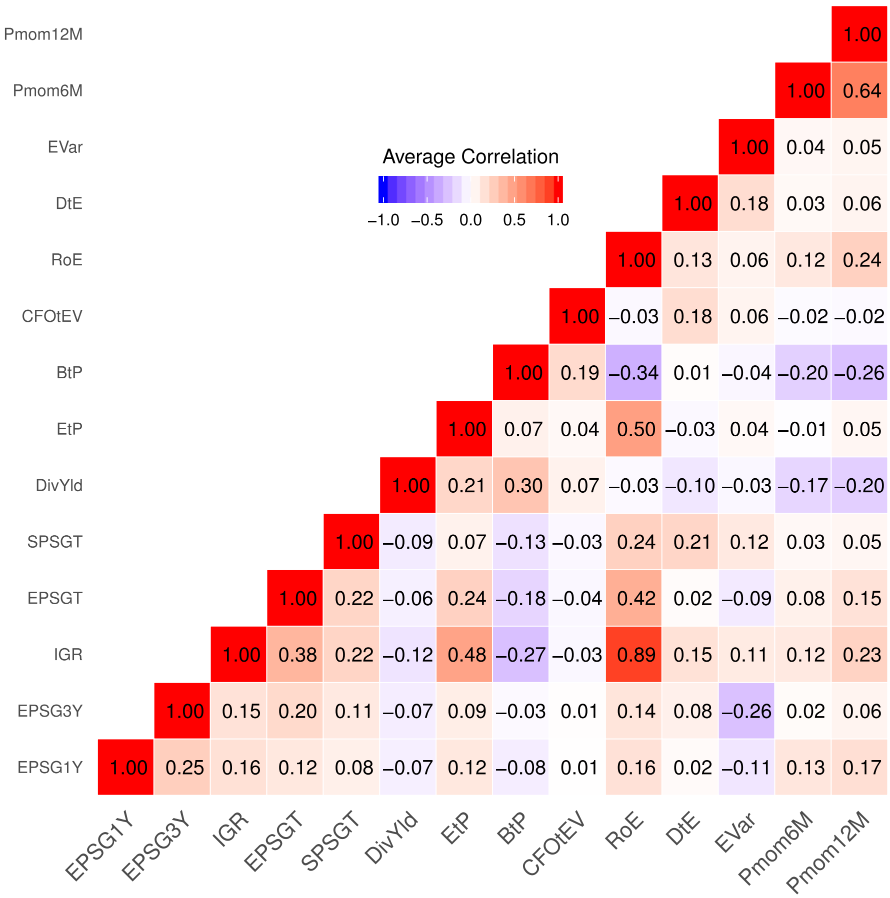

Fig. 3 Average correlations between firm characteristics in the NKY sample

are resampled randomly with replacement from a total number of 180 estimates from the observations used in Table 2.

4.2 Portfolio construction

Within each of 1,000 simulation samples, we simulate $T = 180$ observations of monthly return and firm characteristic observations. Equivalent to our backtest design in Sect. 5 we apply a rolling estimation window over 60 observations with monthly rebalanced LS portfolios.[10] The active portfolio weights from the GK approach are denoted by $a_{\mathbb{A},GK,t} = a_{long,GK,t} - a_{short,GK,t}$, and for the NZ approach by $a_{\mathbb{A},NZ,t} = a_{long,NZ,t} - a_{short,NZ,t}$, where we have used the definitions for the *long* and *short*

[10] Due to symmetric distribution from the normality assumption for the residual returns and signal observations, long-only portfolios lead to equivalent statements.

weight vector from Eqs. (2) and (3). Alpha forecasts for each security $i = 1, \ldots, N$ from the GK method in $\hat{\alpha}_{GK,t}$ are calculated by

$$\hat{\alpha}_{GK,i,t} = \hat{k}\,\hat{C}^{-1}\,z_{i,t-1}, \tag{9}$$

with

$$\hat{k} = \left(\widehat{IC}_1 \cdots \widehat{IC}_M\right), \quad \widehat{IC}_m = \frac{1}{59}\sum_{t=1}^{59} \mathrm{Corr}\left[\theta_t, z_{m,t-1}\right], \quad \hat{C} = \frac{1}{60}\sum_{t=1}^{60} \hat{C}_t.$$

An IC estimate in the $(M \times 1)$-vector $\hat{k}$ is determined by the time series average of the 59 cross-sectional correlations among the lagged firm characteristics with the residual return observations. Residual returns in the $(N \times 1)$-vector θ_t are calculated from time series regressions of security returns against their benchmark returns. The signal correlation matrix $\hat{C}$ is estimated by the time series average of the cross-sectional correlations $\hat{C}_t$. Note, that in the simulations we neglect estimation errors of z-score standardization by assuming that characteristic observations are equal to the standardized z-scores. The naïve alpha forecast of a security i with the NZ method in $\hat{\alpha}_{NZ,t}$ is calculated by[11]

$$\hat{\alpha}_{NZ,i,t} = \sum_{m=1}^{M} z_{i,m,t-1}. \tag{10}$$

4.3 Properties of IC estimates

To measure the capability to estimate the true ICs, we calculate measures for bias and precision. The bias for a signal type m is calculated as the estimates average deviation from the true information coefficient $(\overline{\widehat{IC}_m - IC_m})$. Precision is calculated as the root-mean-squared error (RMSE), the square root of the average squared deviation $((\overline{\widehat{IC}_m - IC_m})^2)^{1/2})$. The related results are shown in Table 3. Here, low RMSE values indicate high precision and vice versa. Averages are calculated for all estimates across the 1,000 simulation samples. The reported bias results are quite low, which indicates that positive and negative estimation errors are canceled out in the long run, as is expected when using ML estimation. Therefore, the average IC estimates $\overline{\widehat{IC}_m}$ are very close to the mean of the true parameters from Table 2, yet the precision results indicate that estimation errors are high.

Due to a higher number of cross-sectional observations, the precision in the SPX and STX samples, which have $N = 500$ and $N = 600$ observations, respectively, can be improved compared to the NKY sample, which has only $N = 225$ observations.

[11] Notice that we have left out $\sigma_{\varepsilon_{i,t}}$ in Eq. (9) or $\frac{1}{M}$ in Eq. (10) because in the cross-sectional setting these are the same for all securities and do not influence the portfolio weights.

Table 3 IC estimation results of the SPX, STX, and NKY samples

Sample	Factor	Characteristic	$\overline{\widehat{\mathrm{IC}}_{\mathrm{m}}}$	Bias	Precision	(> 0) tr. (%)	(> 0) est. (%)	DHR (%)
SPX	Growth	EPSG1Y	0.007	0.000	0.068	54.3	77.0	52.4
		EPSG3Y	0.004	0.000	0.059	51.5	65.5	50.3
		IGR	0.013	0.000	0.075	60.0	87.9	57.5
		EPSGT	0.012	0.000	0.094	54.4	81.8	52.7
		SPSGT	0.003	0.000	0.073	51.7	63.1	50.5
	Value	DivYld	0.009	− 0.001	0.085	55.1	77.3	52.8
		EtP	0.005	0.000	0.089	49.3	65.3	49.9
		BtP	− 0.004	0.000	0.106	46.5	39.0	50.8
		CFOtEV	0.027	− 0.002	0.084	63.7	98.9	63.4
	Quality	RoE	0.015	− 0.001	0.084	62.2	88.8	59.4
		DtE	0.006	0.000	0.060	55.0	73.9	52.3
		EVar	0.006	0.000	0.067	56.2	71.2	52.6
	Momentum	Pmom6M	0.001	0.000	0.114	55.0	53.3	50.3
		Pmom12M	0.007	0.000	0.129	57.2	66.6	52.5
STX	Growth	EPSG1Y	0.025	− 0.002	0.067	64.8	99.6	64.7
		EPSG3Y	0.007	− 0.001	0.056	54.3	80.0	52.5
		IGR	0.010	− 0.001	0.066	58.2	85.9	56.1
		EPSGT	0.025	0.003	0.101	60.5	97.8	60.0
		SPSGT	0.007	0.001	0.076	57.1	73.7	53.2
	Value	DivYld	− 0.002	0.000	0.093	46.1	45.2	50.2
		EtP	0.018	− 0.002	0.081	60.4	95.2	59.5
		BtP	− 0.007	0.001	0.129	48.8	33.2	50.3
		CFOtEV	0.026	− 0.002	0.067	71.6	99.8	71.5

Table 3 continued

Sample	Factor	Characteristic	$\overline{\widehat{\mathrm{IC}}_{\mathrm{m}}}$	Bias	Precision	(> 0) tr. (%)	(> 0) est. (%)	DHR (%)
	Quality	RoE	0.024	− 0.002	0.097	63.4	97.0	62.5
		DtE	0.017	− 0.002	0.079	55.6	94.9	54.9
		EVar	0.002	0.000	0.065	55.2	58.4	50.9
	Momentum	Pmom6M	0.015	− 0.002	0.135	57.3	81.9	54.5
		Pmom12M	0.032	− 0.003	0.151	58.5	96.2	57.8
NKY	Growth	EPSG1Y	0.027	− 0.004	0.086	60.3	98.0	59.9
		EPSG3Y	0.002	0.000	0.090	52.5	55.7	50.4
		IGR	0.006	− 0.001	0.126	51.0	62.8	50.1
		EPSGT	0.010	0.000	0.112	52.5	75.0	51.1
		SPSGT	0.003	0.000	0.102	53.4	57.5	50.6
	Value	DivYld	0.019	− 0.002	0.099	57.1	90.1	55.8
		EtP	0.027	− 0.003	0.110	58.8	95.8	58.1
		BtP	0.027	− 0.003	0.123	58.7	94.5	57.8
		CFOtEV	0.031	− 0.003	0.094	65.1	98.6	64.7
	Quality	RoE	0.005	0.000	0.129	50.3	61.7	50.3
		DtE	0.009	− 0.001	0.120	52.3	70.1	51.1
		EVar	0.023	− 0.003	0.126	64.2	91.0	61.6
	Momentum	Pmom6M	− 0.012	0.001	0.143	46.0	26.7	52.1
		Pmom12M	− 0.007	0.000	0.148	48.9	35.6	50.4

This table shows average IC estimates, measures for bias and precision, the percentages of the true positive ICs ((> 0) tr.), the percentages of estimated positive ICs ((> 0) est.), and the direction hit ratio. These measures are calculated for all IC estimates across the 1000 simulation samples

Nevertheless, the precision is revealed to be low in all three samples. Because of low estimation precision and location parameters that are close to zero, it is important to analyze how often $\widehat{IC}_m$ is able to predict the correct sign of the true IC. Column (> 0) tr. in Table 3 provides the percentages of positive true ICs among all simulation samples. Since the NZ method always assumes positive ICs, these percentages capture the number of IC signs, which were correctly matched using this assumption. Similarly, in column (> 0) est., the percentages of all positive estimated ICs are shown, whereas the direction hit ratio (DHR) depicts the percentages of estimated ICs, which were able to capture the true IC sign. The direction hit ratios indicate that estimating the true sign suffers from low estimation precision, where mean ICs are very close to zero and not significant.

In contrast, significant and relatively high ICs indicate higher direction hit ratios. Since these signal types will have higher estimated mean ICs, signal types with a high DHR will have a higher weight in the alpha forecast. Therefore, even if the average DHR is smaller than in the NZ method, the GK approach can provide an overall better forecast by weighting these signal types higher, that have a better probability of hitting the right direction.

4.4 Actual portfolio alphas and information ratios

Table 4 provides simulation results for average annualized actual portfolio alphas and information ratios. Mean values, standard errors (SE), 0.025 and 0.975 quantiles (Qt.) for the 1000 simulation samples are calculated across the averages of annualized monthly performance results over the evaluation interval of 120 months. Alpha parameters are annualized by multiplying monthly values by 12, whereas standard deviations of monthly residuals in the information ratios are multiplied by the square root of 12.

Realized actual portfolio alphas for $\alpha_{\mathbb{A},GK,t}$ and $\alpha_{\mathbb{A},NZ,t}$ for the active portfolio weights $a_{\mathbb{A},GK,t-1}$, $a_{\mathbb{A},NZ,t-1}$ and true alpha parameters α_t are computed as

$$\alpha_{\mathbb{A},GK,t} = a'_{\mathbb{A},GK,t-1}\,\alpha_t, \tag{11}$$

$$\alpha_{\mathbb{A},NZ,t} = a'_{\mathbb{A},NZ,t-1}\,\alpha_t. \tag{12}$$

Therefore, the realized actual portfolio information ratios $\mathrm{IR}_{\mathbb{A},GK,t}$ and $\mathrm{IR}_{\mathbb{A},NZ,t}$ for $a_{\mathbb{A},GK,t-1}$, $a_{\mathbb{A},NZ,t-1}$ and the true covariance matrix of residuals Ψ are determined by

$$\mathrm{IR}_{\mathbb{A},GK,t} = \frac{\alpha_{\mathbb{A},GK,t}}{\sqrt{a'_{\mathbb{A},GK,t-1}\,\Psi\,a_{\mathbb{A},GK,t-1}}}, \tag{13}$$

$$\mathrm{IR}_{\mathbb{A},NZ,t} = \frac{\alpha_{\mathbb{A},NZ,t}}{\sqrt{a'_{\mathbb{A},NZ,t-1}\,\Psi\,a_{\mathbb{A},NZ,t-1}}}. \tag{14}$$

Table 4 Simulation results from long-short portfolios of the SPX, STX, and NKY samples

Factor		Method	SPX				STX				NKY			
			Mean	SE	0.025-Qt.	0.975-Qt.	Mean	SE	0.025-Qt.	0.975-Qt.	Mean	SE	0.025-Qt.	0.975-Qt.
Multif.	$\overline{\alpha}_{\mathbb{A}}$	NZ	**4.4%**	1.7%	1.1%	7.7%	**8.1%**	1.6%	5.1%	11.5%	**8.4%**	2.5%	3.7%	13.4%
		GK	**5.5%**	1.9%	1.4%	9.3%	**10.7%**	2.6%	6.3%	15.9%	**10.7%**	2.6%	5.8%	15.7%
	$\overline{\mathrm{IR}}_{\mathbb{A}}$	NZ	**1.550**	0.590	0.388	2.702	**2.887**	0.585	1.819	4.107	**1.960**	0.578	0.855	3.123
		GK	**1.923**	0.678	0.508	3.246	**3.847**	0.927	2.304	5.748	**2.501**	0.609	1.363	3.662
Growth	$\overline{\alpha}_{\mathbb{A}}$	NZ	0.9%	1.1%	– 1.3%	3.2%	1.2%	1.2%	– 0.9%	3.5%	2.5%	1.9%	– 1.4%	6.3%
		GK	1.9%	2.0%	– 1.7%	5.9%	**7.1%**	2.3%	2.6%	11.2%	4.6%	2.4%	– 0.3%	8.8%
	$\overline{\mathrm{IR}}_{\mathbb{A}}$	NZ	0.323	0.403	– 0.496	1.125	0.423	0.420	– 0.333	1.238	0.573	0.455	– 0.325	1.469
		GK	0.668	0.694	– 0.632	2.056	**2.550**	0.810	0.911	4.049	1.084	0.552	– 0.062	2.084
Value	$\overline{\alpha}_{\mathbb{A}}$	NZ	3.2%	2.1%	– 0.9%	7.2%	3.4%	2.3%	– 1.2%	7.7%	**10.3%**	2.4%	5.7%	14.9%
		GK	**5.5%**	1.7%	1.9%	8.8%	**7.3%**	1.7%	3.9%	10.6%	**9.3%**	2.7%	4.0%	14.2%
	$\overline{\mathrm{IR}}_{\mathbb{A}}$	NZ	1.132	0.725	– 0.331	2.504	1.204	0.820	– 0.446	2.790	**2.412**	0.570	1.305	3.480
		GK	**1.934**	0.612	0.651	3.064	**2.609**	0.617	1.413	3.775	**2.178**	0.635	0.919	3.314
Quality	$\overline{\alpha}_{\mathbb{A}}$	NZ	3.3%	1.8%	– 0.3%	6.6%	**6.0%**	2.1%	2.0%	10.4%	4.8%	2.9%	– 0.9%	10.4%
		GK	2.4%	2.1%	– 1.7%	6.6%	**6.4%**	2.6%	0.6%	11.6%	3.6%	3.3%	– 2.7%	9.9%
	$\overline{\mathrm{IR}}_{\mathbb{A}}$	NZ	1.143	0.614	– 0.097	2.293	**2.159**	0.745	0.721	3.715	1.118	0.683	– 0.184	2.453
		GK	0.850	0.753	– 0.566	2.303	**2.310**	0.945	0.252	4.181	0.853	0.763	– 0.646	2.317
Momentum	$\overline{\alpha}_{\mathbb{A}}$	NZ	1.0%	2.5%	– 3.9%	6.1%	**6.4%**	3.0%	0.1%	12.5%	– 2.2%	3.0%	– 7.9%	3.7%
		GK	0.7%	2.3%	– 3.8%	5.5%	**7.0%**	3.3%	0.5%	13.2%	1.1%	2.9%	– 4.3%	6.9%
	$\overline{\mathrm{IR}}_{\mathbb{A}}$	NZ	0.346	0.873	– 1.333	2.108	**2.290**	1.088	0.076	4.451	– 0.504	0.701	– 1.837	0.861
		GK	0.258	0.815	– 1.326	1.947	**2.505**	1.182	0.205	4.703	0.250	0.681	– 0.991	1.616

Average alphas and information ratios are calculated across the 120 monthly results in each simulation sample. Mean, standard errors and quantiles are calculated across the 1,000 simulation samples. Significant mean values, where the confidence bounds are not crossing the zero value, are marked as bold

Due to the application of the true alphas and covariance matrices with the estimated portfolio weights, we receive performance values, which represent the actual instead of the estimated performance results. Since our assumption for Ψ is based on shrinkage estimation the application of Ψ in the volatility calculations leads to lower quantities than the estimates from historical return data in common backtest studies. Therefore, the actual information ratios shown in the simulation should be higher than the estimates in the backtest study.

Besides the multifactor portfolio (Multif.), where all 14 firm characteristics are assumed to be known by the investor, we additionally evaluate the performance for the single-factor portfolios, where only the subset of the conditioning characteristics are applied.

In the STX and NKY data samples, the highest performance results are achieved by the multifactor portfolios. In the SPX sample, the multifactor and the value factor belong to the best performing portfolios, although for the SPX sample, only a small subset of characteristics indicate statistically significant mean ICs. Consistent with the number of characteristics with significant mean ICs, the multifactor results in the STX and NKY samples show almost twice the mean alpha results compared to the SPX sample.

In the comparison between the GK and NZ approaches, the results reveal that, in the multifactor portfolios, the GK and NZ approaches are both able to significantly outperform the benchmark portfolio. In all three samples, the results of the GK approach are higher than in the NZ approach, although both methods perform similarly. In the single-factor portfolios, the performance results indicate that, in factors with a small number of significant mean ICs, neither approach is able to significantly outperform the benchmark portfolio. The growth, quality or momentum factors of the SPX and NKY samples are good examples of these findings. Even if there exists one characteristic with a significant mean IC, estimation errors from the other characteristics can outweigh the benefits of IC application. This is especially the case for factors with many redundant signals, such as the growth portfolios of the SPX and NKY samples.

Besides the number of significant characteristics, the level of the IC also has a performance impact. Characteristics such as CFOtEV, with high mean IC values of approximately 3.0 %, lead to lower estimation errors on average and attain a higher weight compared to characteristics with low mean ICs. Therefore, in the SPX sample, the GK value portfolio is able to achieve significant positive performance results, although it has more redundant characteristics than the quality factor. Furthermore, in factors with a high IC dispersion between the signals, such as the growth and value factors in the STX sample, the GK method is able to incorporate this dispersion. This leads to better average performance results compared to the NZ method, which is not able to beat the benchmark portfolio. In contrast, if the IC parameters are all positive and the IC dispersion is low, as in the value factor in the NKY sample, the NZ method achieves a better portfolio performance than the GK approach, though the difference is not large. As the NZ method implicitly assumes positive ICs, the GK method should be advantageous for factors with negative mean ICs. The results from the momentum factor in the NKY sample support this assumption. The NZ method indicates negative performance, whereas the GK method shows positive alphas and information ratios; however, due to low and not significant mean ICs, the performance results are not significant.

5 Out-of-sample backtest

In the following sections, the backtest and its results will be described and analyzed. The framework of the backtest is chosen to resemble the realistic investing behavior of an institutional investor, whose main aim is to outperform an underlying cap-weighted benchmark portfolio. This includes the consideration of long-only constraints, rebalancing costs, commonly used rebalancing frequencies and representative data sets. We first analyze the portfolios' security allocation performance by confirming that all index constituents are selected. In addition, we analyze selection effects for the multifactor portfolio by separating the index constituents into quantile portfolios.

Consistent with our simulation study, we use a rolling window of five years, which corresponds to 60 monthly observations. The out-of-sample evaluation interval starts in 2007. Securities are weighted proportionally to their alpha forecasts according to Eqs. (2) and (3), where forecasts are conducted as in formulas (9) and (10). However, in contrast to the simulation, ratio observations are standardized.

5.1 Rebalancing costs

We assume that the signals are available at the end of each month. Thus, the portfolio is rebalanced at the beginning of the next month. To remain as close as possible to an investable portfolio, we take into account trading costs caused by the portfolio rebalancing. To quantify the cost-relevant volume the portfolio turnover rate ($\text{PTR}_t = \sum_{i=1}^{N} |a_{i,t-1} - a_{i,t}|$) is used. The PTR determines the percentage of the portfolio that causes trading costs.

Exemplary, in this context a PTR of 30 % signifies that we have to sell 15 % of the old portfolio. Subsequently, we are obligated to reinvest the incoming liquidity and, as a result, trading costs are incurred for 30 % of the portfolio. In the backtest we assume the amount of costs at ten basis points per volume traded. This value comprises the trading commission, which should not exceed the amount of five basis points, and a trading tax, which increases depending on the specific country's regulations. Normally, the costs incurred in connection with taxes clearly exceed the costs related to trading. However, as only a few countries impose a tax on financial transactions, average tax costs of five basis points appear to be realistic. Furthermore, Frazzini et al. (2018) measure real-world trading costs for a broad international data set of executed live trades and confirm our assumptions with empirical results for a recent sample from 2006 to 2016. The calculated costs are demarcated on the day they arise. Consequently, these costs lead to a direct reduction of the return on the rebalancing day.

5.2 Perfect foresight results

To evaluate the maximum potential performance contribution of incorporating ICs and signal correlations, we compare the perfect foresight portfolio results of the NZ and GK methods. This means that, on each rebalancing date, we assume to know the true realized ICs and signal correlations beforehand (GK PF III). In addition to the perfect foresight case, we also investigate a portfolio, in which we replace the true signal

Table 5 Perfect foresight results

	Alpha Est. Method	Alpha			Information ratio		
		SPX (%)	STX (%)	NKY (%)	SPX	STX	NKY
Long-only	NZ	1.5	3.4	3.0	0.306	0.745	0.608
	GK PF I	0.5	2.5	2.3	0.117	0.539	0.453
	GK PF II	24.0	25.1	31.9	2.374	3.308	3.496
	GK PF III	28.7	28.3	37.7	2.661	3.577	3.833
Long-short	NZ	4.3	6.9	4.1	0.688	1.608	0.442
	GK PF I	2.4	4.7	3.3	0.428	1.178	0.471
	GK PF II	51.3	52.6	62.2	4.539	5.409	6.525
	GK PF III	59.5	59.1	72.8	4.986	5.935	6.729

This table presents alphas and information ratios of the perfect foresight (GK PF) as well as of the naïve approach (NZ). GK PF III represents the perfect foresight case, where we assume to know the true realized ICs and the true signal correlations. GK PF I shows the case, where ICs are constant and the true signal correlations are known, whereas GK PF II shows the results where we pretend to know the true ICs and replace the true signal correlation matrix with the unit matrix

correlation matrix with the unit matrix and only pretend to know the true ICs (GK PF II) as well as a portfolio with constant ICs and the true signal correlation matrix (GK PF I). This examination enables us to evaluate the difference in the impact of ICs and the signal correlations on the alpha and information ratio.

In Table 5, annualized information ratios and the alpha intercept coefficients from the regressions of the monthly portfolio returns against their according index returns are reported. As illustrated for the multifactor portfolio, if we know the true ICs and the signal correlation matrix, it would be possible to increase the portfolio's performance considerably. The attainable alphas with perfect foresight in GK PF III are between 8.3 and 19.1 times higher than the alphas achieved by using the NZ approach. In addition to the increasing alphas, we observed a corresponding increase in the information ratios. Here, the values increased by a factor of at least 3.7 up to a factor of 15.9. The comparison of the GK PF I and GK PF II results reveals, that the consideration of the signal correlation matrix can increase portfolio performance, but to a substantially lesser extent than ICs. It can be stated that the main performance share is driven by IC estimation. Additionally, the low precision of IC estimation from Sect. 4.3 and the abnormal high-performance quantities from the perfect foresight results indicate potential benefits by reducing IC estimation errors, which leaves room for possible further research.

Moreover, the perfect foresight results provide interesting findings regarding the turnover of the strategies. While there are no substantial differences between the NZ, GK and GK PF I portfolios which are showing an average turnover of 21.7%, the GK PF II and GK PF III portfolios are showing a significant increased average turnover of 65.4%. Therefore, the trading costs of the portfolios, where we assume to know the true realized ICs, are three times higher than for the other portfolios. But in this case, a high turnover is not surprising, as the true realized ICs are fluctuating randomly across a common mean. Consequently, if we assume to know the real ICs the portfolio weights are subject to strong changes. However, as our calculations take into account trading costs, the results show that both approaches are able to overcompensate the higher costs by significantly higher performance results.

5.3 Backtest results

To examine their performance in comparison to their associated cap-weighted benchmark portfolios, the annualized alpha parameters, information ratios and Sharpe ratios for each factor strategy are reported in Table 6. In addition to the factor portfolios, an equally weighted $1/N$ portfolio has been evaluated. We use the time series of monthly portfolio returns to compute the out-of-sample portfolio alphas, which are estimated from a regression of the portfolio returns against the associated benchmark returns. To determine the strategies' information ratios, we divide the portfolios' alpha parameters by the tracking errors, calculated from the standard deviations of the regression residuals. The Sharpe ratios are computed ex-post using the portfolios' realized returns. The three performance measures lead to consistent results for most cases when comparing the GK and NZ approaches. One exception can be found in the LS STX multifactor portfolios. Here, the volatility of the NZ portfolio return is exceptionally low. Comparing to the GK approach, this low volatility leads to a higher information ratio, although the GK approach achieves a higher alpha result.

Equivalent to the amount of significant mean ICs in Table 2, the highest amount of portfolios with significant alpha results can be observed for the STX sample, followed by the NKY and the SPX samples. Moreover, the $1/N$ portfolios, which do not take into account any firm characteristics, fail to significantly outperform the benchmark portfolios. For the most part, the multifactor performance appears to be beneficial in comparison to the single-factor portfolios. For the LS multifactor portfolios, it can be noted that the GK and NZ approach could both achieve significant alpha parameters in almost all samples.

However, in the LO portfolios of the SPX sample, both strategies fail to outperform the benchmark, which is consistent with the low amount of significant mean ICs. Even in the LS multifactor portfolios, the results of the SPX sample appear to be weak. Interestingly, in the SPX multifactor portfolios, the NZ approach reveals to have a better performance than the GK method. This supports the result, that the GK approach suffers from estimation errors caused by redundant characteristics, even though there is no significant difference between both approaches. In contrast to the SPX sample, in the STX sample, the LO and LS multifactor portfolios outperform the benchmark. Here, the performance of the GK portfolios is higher than in the NZ portfolios, but since most of the ICs are positive and the overall IC dispersion is not high there is no significant difference. Because the IC structure within the NKY multifactor portfolio indicates a high informational content concentration among the value characteristics, we can indicate a high IC dispersion in the multifactor portfolio. Consequently, in the LS multifactor portfolio, the GK approach is able to outperform the benchmark, while the NZ approach fails to do so.[12] Consistent with the multifactor results, in the

[12] To visualize the potential performance benefits, in Figs. 4 and 5 in the appendix, the cumulative portfolio returns for the GK and NZ multifactor portfolios for the LO and LS portfolios are depicted.

Table 6 Backtest results of the SPX, STX, and NKY samples

Factor	Method		Alpha			Information Ratio			Sharpe Ratio		
			SPX	STX	NKY	SPX	STX	NKY	SPX	STX	NKY
	$1/N$		0.4%	0.9%	0.8%	0.100	0.213	0.170	0.470	0.268	0.257
Multif.	LO	NZ	1.5%	3.4%**	3.0%*	0.306	0.745	0.608	0.593	0.446	0.367
		GK	0.5%	4.3%**	4.9%**	0.108	0.794	0.711	0.541	0.496	0.438
	LS	NZ	4.3%**	6.9%***	4.1%	0.688	1.608	0.442	0.512	1.497	0.417
		GK	3.3%*	8.0%***	8.7%***	0.509	1.007	1.346	0.266	0.857	1.343
Growth	LO	NZ	– 0.5%	1.6%	1.6%	– 0.100	0.365	0.293	0.489	0.346	0.303
		GK	0.2%	4.0%**	0.2%	0.033	0.696	0.038	0.520	0.473	0.239
	LS	NZ	0.4%	2.2%	1.5%	0.084	0.521	0.266	0.120	0.453	0.254
		GK	2.9%	6.0%**	0.7%	0.420	0.730	0.101	0.196	0.660	0.076
Value	LO	NZ	2.2%	0.8%	6.7%***	0.207	0.090	0.922	0.567	0.275	0.516
		GK	2.0%	3.4%***	5.9%**	0.392	0.634	0.817	0.619	0.436	0.477
	LS	NZ	2.4%	– 0.7%	9.9%***	0.214	– 0.067	1.300	0.405	0.096	1.321
		GK	4.2%*	5.6%***	8.1%***	0.629	0.868	1.125	0.375	0.730	1.149
Quality	LO	NZ	0.8%	3.8%***	0.3%	0.271	0.852	0.072	0.572	0.487	0.246
		GK	– 0.5%	3.8%**	– 1.2%	–0.105	0.770	– 0.285	0.483	0.479	0.177
	LS	NZ	4.3%*	6.3%***	– 0.6%	0.592	0.911	– 0.0538	0.116	0.586	– 0.092
		GK	– 0.6%	7.5%***	– 1.5%	–0.119	1.017	– 0.162	–0.138	0.786	– 0.207
Momentum	LO	NZ	– 0.6%	2.6%	– 2.7%	–0.152	0.508	– 0.442	0.472	0.422	0.103
		GK	– 0.7%	1.9%	– 1.3%	–0.166	0.340	– 0.171	0.471	0.368	0.171
	LS	NZ	0.7%	5.0%	– 5.0%	0.069	0.475	– 0.428	– 0.175	0.287	– 0.453
		GK	– 0.1%	2.8%	– 1.3%	–0.015	0.285	– 0.141	– 0.146	0.165	– 0.123

Significant alpha parameters are marked with asterisks. *,**,***represent the significance of the test, i.e., the rejection of the null hypothesis, with a level of probability of 10%, 5%, and 1%, respectively

Table 7 Alpha parameters of the GK portfolios against the NZ portfolios

Factor	Portfolio	Alpha		
		SPX (%)	STX (%)	NKY (%)
Multifactor	LO	− 0.7	1.0	1.8
	LS	0.5	1.9	7.9 ∗ ∗∗
Growth	LO	0.8	2.5	− 1.4
	LS	1.3	6.8 ∗ ∗∗	− 0.2
Value	LO	1.5	3.2∗	− 0.9
	LS	4.0∗	5.4 ∗ ∗∗	− 0.7
Quality	LO	− 1.5	− 0.1	− 1.5 ∗ ∗
	LS	− 0.8	2.1∗	− 1.2
Momentum	LO	0.2	− 0.7	1.8
	LS	− 0.1	− 0.9	− 3.5

Significant alpha parameters are marked with asterisks. *,**,***represent the significance of the test, i.e., the rejection of the null hypothesis, with a level of probability of 10%, 5%, and 1%, respectively

single-factor portfolios, we can observe that the IC structure determines the potential performance benefits of the GK approach. For instance, the high IC dispersion structure of the value and growth factors within the STX sample favors the GK method, as this approach is able to incorporate the dispersion and consequently achieves significant performance results. However, the NZ approach fails to outperform the benchmark in these factors. A special case is represented by the NKY value portfolio, where all mean ICs are significantly positive and have low dispersion. This type of IC structure represents a situation in which an investor has a well-defined factor model and informational content is approximately equal among all characteristics. These conditions are particularly suitable for the NZ approach, which in the LS value portfolio reaches the highest performance of all investigated factor portfolios. However, it is remarkable that the GK approach is also able to achieve a similarly high performance. By comparison, the quality factor of the SPX and NKY samples display factor portfolios whose characteristics show predominantly insignificant mean ICs. Here, the GK approach suffers from IC estimation errors, which leads to negative performance results. In contrast, the NZ method significantly outperforms the SPX in the LS quality portfolio, which demonstrates that averaging over the applied firm characteristics may benefit from lower estimation errors. Interestingly, there is no significant outperformance in the momentum factor portfolios in any of the three samples. Our results are in line with findings of Lewellen (2015, p. 12), who argues that due to the financial crisis several months in 2009 resulted in very poor performance of momentum strategies which reduces the out-of-sample performance of return forecasts and is especially apparent for larger stocks.

The annualized alpha results in Table 7 confirm the findings from Table 6. Alphas are calculated from the intercept coefficient based on regressions of the GK portfolio returns against the NZ portfolio returns. In the STX value and growth factors, which show high IC dispersion, the GK approach outperforms the NZ method. Additionally, the GK method demonstrates a significant alpha for the LS NKY multifactor portfolio.

Table 8 Alpha and information ratio results of the quantile portfolios

Method	Interval	Alpha			Information ratio		
		SPX (%)	STX (%)	NKY (%)	SPX	STX	NKY
NZ	0.8–1	1.4**	3.3*	3.6*	0.284	0.669	0.627
	0.6–0.8	1.2	1.8	1.0	0.273	0.387	0.201
	0.4–0.6	− 1.3	0.7	− 3.4**	− 0.283	0.167	− 0.673
	0.2–0.4	− 0.9***	− 0.8	1.1	− 0.161	−0.183	0.199
	0–0.2	− 3.9	− 4.5***	− 3.4	− 0.593	− 0.903	− 0.381
GK	0.8– 1	0.4	4.5**	5.7**	0.093	0.751	0.769
	0.6–0.8	− 0.9	2.9**	1.2**	− 0.198	0.740	0.196
	0.4–0.6	0.4	0.9	1.0	0.107	0.212	0.196
	0.2–0.4	− 0.5	− 2.8**	− 4.7**	− 0.089	− 0.650	− 0.826
	0–0.2	− 3.1	− 5.2**	− 3.9*	− 0.445	− 0.686	− 0.618

Significant alpha parameters are marked with asterisks. *,**,***represent the significance of the test, i.e., the rejection of the null hypothesis, with a level of probability of 10%, 5%, and 1%, respectively

However, the GK approach achieves a significantly negative alpha for the LO quality factor in the NKY sample. Here, the only signal with significant mean IC is EVar with a low mean value of 1.8 %, whereas the other two signals, RoE and DtE with mean ICs of 0.7 %, appear to be redundant.

5.4 Quantile portfolios

In the previous section, we investigated allocation results, by selecting all index constituents. In theory and practice, there is a tendency to build subportfolios where the long portfolio contains the top quantile and the short portfolio contains the bottom quantile of the securities. To investigate the impact of this selection approach, we separate the total amount of index securities according to their alpha forecasts into five quantile portfolios. As securities can be divided into two groups, one with positive and the other with negative alpha forecasts, the upper two quantile portfolios should achieve a positive alpha and the lower two should turn out to be negative. To demonstrate that portfolio alphas decrease in the lower quantile portfolios, we construct long-only portfolios in which the forecasted alphas have to be transformed to receive only positive alpha values:

$$\alpha\text{-score}_i = \begin{cases} 1 + \hat{\alpha}_i & \text{if} \quad \hat{\alpha}_i \geq 0, \\ \left(1 - \hat{\alpha}_i\right)^{-1} & \text{if} \quad \hat{\alpha}_i < 0. \end{cases} \tag{15}$$

Securities with alpha-scores greater than one represent positive alpha forecasts, which are located in the upper quantiles. Securities with an alpha-score between zero and one represent negative alphas, which are assigned to lower quantile portfolios. The alpha results in Table 8 are calculated using regressions of the quantile portfolio returns against the corresponding index returns. Again, the results appear to be beneficial for the GK approach in the well-defined factor model within the STX sample. The exterior, as well as the second and fourth quantile portfolios, show significant outperformance, while in the NZ portfolios, only the exterior portfolios attain significant results. This indicates that the GK method is more appropriate for selecting securities in the second and the fourth quantile than the NZ approach, which would be beneficial for portfolios with a high amount of constituents. A similar outcome can be observed in the NKY sample, although the middle quantile in the NZ approach appears to be significantly negative. This underlines the challenging task for the NZ approach to separate between well and poorly performing securities if the multifactor model exhibits a high concentration of informational content in a single factor. Consistent with our previous results, the GK performance for the SPX sample shows no significant results, while the NZ approach appears to beat the market to at least a small extent for the highest and second highest quantiles.

6 Conclusion

We apply factor investing strategies motivated by investing in the optimal orthogonal portfolio, where portfolio weights are proportional to the securities' alpha parameters. Performance results from simulations and out-of-sample backtests indicate that factor portfolios, which take into account multiple firm characteristics, are able to significantly outperform their corresponding cap-weighted benchmark portfolios. In comparison to stock screening with naïve weighted z-scores, we have investigated performance gains caused by incorporating a linear alpha forecasting model, which allows us to take into account differences in correlations and informational content between different types of firm characteristics. Here, we focus our analysis on ICs, which could be identified as the main performance drivers. Furthermore, our study illustrates the importance of ICs as decision variables within the discrimination between firm characteristics. Simulation and backtest performance results show that the disadvantage of estimation errors in the ICs can be outweighed by the benefits of incorporating these into the securities' alpha forecasts if the applied factors are well-defined and show a high dispersion between the informational content of different firm characteristics. This is the case for the investigated multifactor portfolio in the NKY sample as well as the value and growth factor portfolios in the STX sample, where the proposed strategy outperforms the naïve approach. However, the GK approach underperforms in factor models with low IC dispersion, whereas the NZ approach is able to benefit from lower estimation errors. Our results should motivate practitioners to deviate from equally weighted scoring approaches when they believe in their underlying factor model and want to discriminate between characteristics with high IC dispersion. Due to limited space, we apply a very simple linear forecasting model. Thus, further exploring the implications of time-varying predictability would be a worthwhile avenue for continued research.

Acknowledgements We would like to thank the editor Markus Schmid and the anonymous referees for their constructive recommendations, which helped to improve the quality of this paper.

Appendix

See Tables 9, 10, 11 and Figs. 4, 5.

Table 9 Annualized mean returns and standard deviations of the benchmark returns and risk-free rates

	$R_{\mathbb{B}}$			R_f		
	SPX (%)	STX (%)	NKY (%)	SPX (%)	STX (%)	NKY (%)
Mean	8.2	4.8	5.1	0.6	0.9	0.2
Sd.	15.1	15.3	20.8	1.4	1.5	0.2

Table 10 Descriptive statistics of the firm characteristics (in %)

	Characteristic	Min.	0.25-Qt.	Median	Mean	0.75-Qt	Max.
SPX	EPSG1Y	− 75.0	−11.9	7.4	8.0	25.0	150.5
	EPSG3Y	− 33.6	4.2	14.9	22.6	31.4	269.2
	IGR	− 145.2	4.6	10.3	10.5	16.4	48.6
	EPSGT	− 43.6	2.2	8.6	8.4	15.3	55.9
	SPSGT	− 12.4	1.4	5.6	6.1	10.2	29.5
	DivYld	0.0	0.6	1.6	1.8	2.6	14.4
	EtP	1.1	5.0	6.1	5.6	7.3	12.2
	BtP	9.6	28.0	40.0	35.5	55.8	131.7
	CFOtEV	− 0.9	5.9	7.8	8.5	10.6	24.7
	RoE	− 34.1	9.6	15.3	16.8	22.0	116.9
	DtE	0.0	35.6	57.9	72.4	97.2	922.3
	EVar	10.9	29.7	46.3	56.0	71.3	297.4
	Pmom6M	− 77.0	−21.6	17.6	18.7	57.2	127.2
	Pmom12M	− 90.0	−27.5	29.5	38.5	98.3	224.8
STX	EPSG1Y	− 149.3	−18.2	6.3	9.6	30.4	235.2
	EPSG3Y	− 34.8	1.4	14.3	24.9	36.4	298.4
	IGR	− 71.6	1.7	7.0	7.3	13.3	44.1
	EPSGT	− 53.4	−5.2	6.7	5.1	16.6	58.0
	SPSGT	− 14.2	0.5	4.9	5.2	9.6	27.7
	DivYld	0.0	2.0	3.0	3.3	4.3	11.7
	EtP	2.6	4.8	6.1	5.9	7.8	19.8
	BtP	14.6	33.3	48.0	43.0	72.8	157.4
	CFOtEV	− 13.7	5.2	7.4	8.4	10.6	44.8
	RoE	− 11.9	8.4	13.9	15.3	20.7	58.5
	DtE	1.7	44.3	68.3	110.5	107.4	802.7
	EVar	14.8	43.5	79.4	103.8	152.1	394.3
	Pmom6M	− 102.9	−25.0	13.2	12.4	50.5	122.6
	Pmom12M	− 124.2	−27.5	31.2	33.9	92.5	219.5

Table 10 continued

	Characteristic	Min.	0.25-Qt.	Median	Mean	0.75-Qt	Max.
NKY	EPSG1Y	− 82.6	−24.3	2.4	6.0	32.3	122.6
	EPSG3Y	− 33.1	1.5	15.8	20.5	35.1	248.5
	IGR	− 48.8	1.2	4.4	3.6	7.3	32.7
	EPSGT	− 63.0	−13.8	5.4	3.6	20.8	66.2
	SPSGT	− 13.7	−2.0	1.7	1.9	5.7	30.6
	DivYld	0.1	1.2	1.7	1.9	2.4	9.6
	EtP	1.7	4.2	5.5	5.1	7.2	15.7
	BtP	11.7	61.5	85.8	73.0	114.5	232.7
	CFOtEV	− 12.9	5.8	8.9	9.4	12.2	87.5
	RoE	− 145.3	3.5	6.8	6.0	10.3	39.4
	DtE	0.1	40.2	73.2	92.8	117.6	795.9
	EVar	19.3	58.5	87.1	120.3	151.2	724.7
	Pmom6M	− 86.4	−32.6	1.4	6.0	40.8	143.1
	Pmom12M	− 102.0	−45.0	1.2	14.6	64.1	254.8

This table shows descriptive statistics of the firm characteristics for all firms in the SPX, STX and NKY samples. These are calculated across all assets in the whole time horizon

Table 11 Sector mean characteristics (in %)

	Characteristic	All	10	15	20	25	30	35	40	45	50	55	60
SPX	EPSG1Y	8.0	11.4	5.5	9.3	7.1	8.1	13.3	1.9	10.6	3.1	6.4	8.0
	EPSG3Y	22.6	37.7	29.0	18.0	18.6	13.4	24.1	18.3	24.5	16.4	17.1	49.3
	IGR	10.5	11.0	10.7	14.3	13.6	12.9	13.5	7.1	12.8	− 0.5	3.6	− 9.2
	EPSGT	8.4	10.6	9.3	11.2	9.1	6.1	10.8	4.6	11.0	2.9	3.6	3.9
	SPSGT	6.1	9.0	4.6	5.6	5.0	4.9	8.8	4.7	7.9	8.0	2.6	7.8
	DivYld	1.8	1.2	2.1	1.9	1.3	2.5	0.6	2.3	0.7	3.9	4.0	3.8
	EtP	5.6	6.3	5.8	6.1	5.9	5.9	5.5	7.6	5.0	6.1	6.7	2.2
	BtP	35.5	49.0	33.2	32.1	31.1	22.8	30.7	75.9	30.1	48.2	62.3	36.6
	CFOtEV	8.5	11.9	7.9	8.4	8.5	6.8	7.1	9.6	9.2	11.9	8.6	3.7
	RoE	16.8	13.9	17.7	20.5	19.0	27.4	17.1	10.4	16.8	16.0	10.8	10.3
	DtE	72.4	47.6	73.3	66.9	69.0	107.4	49.8	101.2	20.8	187.8	127.0	132.3
	EVar	56.0	86.8	79.2	37.6	50.0	31.7	46.0	52.6	64.9	74.6	51.7	101.3
	Pmom6M	18.7	15.1	20.5	21.2	15.3	28.2	24.0	10.7	16.3	14.4	26.5	21.1
	Pmom12M	38.5	27.7	41.3	43.3	36.2	58.2	46.4	25.5	27.2	29.6	56.5	47.8
STX	EPSG1Y	9.6	3.4	11.9	11.1	13.2	9.4	10.2	5.4	17.5	1.8	5.3	14.5
	EPSG3Y	24.9	28.7	27.2	25.3	21.3	16.9	17.3	23.8	29.8	17.8	24.2	62.2
	IGR	7.3	7.6	6.1	8.8	9.8	8.2	10.6	4.8	10.5	− 1.1	4.1	6.9
	EPSGT	5.1	2.4	3.2	6.9	6.6	7.1	9.3	1.3	13.5	1.6	− 0.9	7.4
	SPSGT	5.2	7.0	4.6	5.4	4.5	4.6	7.8	4.4	9.1	1.1	6.2	4.7
	DivYld	3.3	2.7	2.7	3.0	3.3	2.6	1.5	4.4	1.7	5.7	4.9	4.3
	EtP	5.9	6.2	5.5	5.6	5.7	5.1	4.2	8.0	4.3	6.2	6.6	8.3
	BtP	43.0	51.3	52.5	34.5	39.3	32.6	27.4	85.2	30.1	35.1	52.5	96.7
	CFOtEV	8.4	10.3	9.9	7.8	8.6	7.0	6.4	10.0	6.6	12.4	8.8	2.3

Table 11 continued

	Characteristic	All	10	15	20	25	30	35	40	45	50	55	60
	RoE	15.3	14.5	12.4	18.7	16.8	18.7	18.2	10.5	16.5	19.2	14.4	9.9
	DtE	110.5	54.8	53.7	77.4	68.8	69.5	57.7	268.7	33.5	136.1	136.2	84.2
	EVar	103.8	116.4	121.3	91.0	102.3	58.5	61.4	125.5	94.9	107.6	98.2	189.9
	Pmom6M	12.4	5.6	13.8	15.4	13.8	21.6	19.4	2.7	17.1	12.3	10.0	13.7
	Pmom12M	33.9	12.6	32.3	39.0	41.3	53.7	48.0	17.0	41.7	34.1	28.3	37.3
NKY	EPSG1Y	6.0	0.2	0.4	6.4	12.1	10.3	4.1	1.4	4.4	12.1	5.8	15.9
	EPSG3Y	20.5	26.6	17.9	20.4	29.9	16.7	14.7	22.0	11.2	18.6	29.3	35.8
	IGR	3.6	1.6	3.2	4.6	2.5	3.7	4.0	3.8	1.6	10.1	1.1	5.0
	EPSGT	3.6	− 9.3	−1.3	5.8	5.5	5.0	1.6	5.6	2.4	8.9	− 6.6	6.1
	SPSGT	1.9	2.5	1.6	2.2	1.2	1.5	3.3	2.6	0.3	4.4	3.3	3.7
	DivYld	1.9	2.5	2.0	1.8	1.7	1.8	2.2	2.2	1.7	2.1	3.1	1.3
	EtP	5.1	8.5	5.7	5.4	5.0	4.1	4.0	6.8	4.1	7.1	6.0	4.0
	BtP	73.0	99.1	95.2	75.2	64.4	66.8	47.9	95.4	64.3	50.5	92.0	64.8
	CFOtEV	9.4	10.0	10.4	8.2	9.9	8.5	7.0	11.4	11.1	16.0	10.5	4.2
	RoE	6.0	4.0	5.6	7.4	2.7	6.3	8.2	6.2	4.9	13.6	3.1	6.8
	DtE	92.8	73.1	88.8	94.8	60.7	60.1	19.0	228.4	37.4	64.3	164.0	199.8
	EVar	120.3	227.9	123.2	105.0	138.7	92.9	88.1	118.9	160.8	69.5	148.0	99.0
	Pmom6M	6.0	− 3.0	3.2	7.8	5.1	16.2	12.7	− 2.2	3.8	17.8	− 1.5	6.5
	Pmom12M	14.6	− 5.6	7.1	17.3	15.2	34.1	31.8	− 0.8	8.9	40.4	2.7	19.4

This table shows the averages of the firm characteristics for all firms in the SPX, STX and NKY samples across all assets in the whole time horizon. Additionally, the averages are shown for the industry groups separated by GICS sector codes: 10: “Energy”, 15: “Materials”, 20: “Industrials”, 25: “Consumer Discretionary”, 30: “Consumer Staples”, 35: “Health Care”, 40: “Financials”, 45: “Information Technology”, 50: “Telecommunication Services”, 55: “Utilities”, 60: “Real Estate”. Averages are calculated as the grand mean over all observations

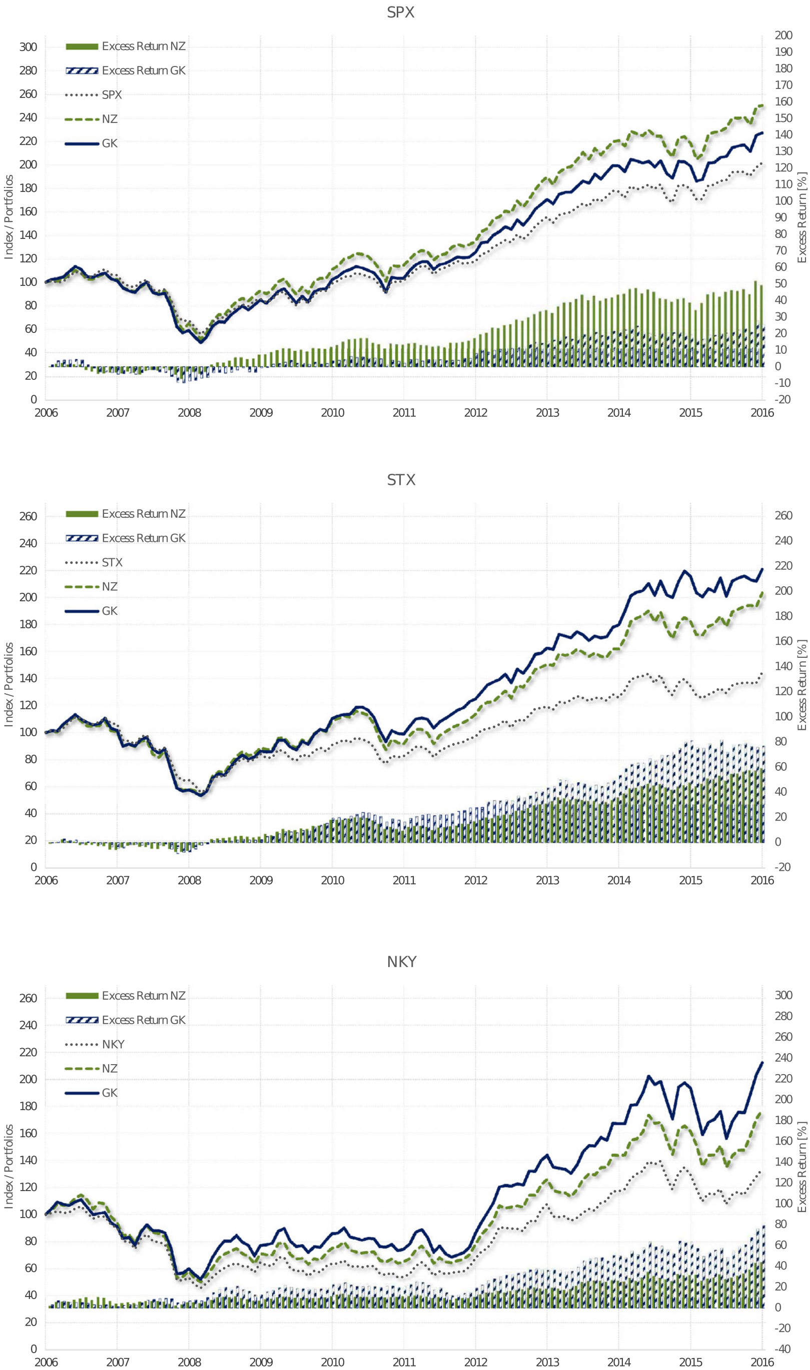

Fig. 4 Cumulative performance results of long-only multifactor portfolios. Cumulative portfolio returns appear on the left axis and excess returns in comparison to the index returns are shown as vertical bars with their corresponding values on the right axis

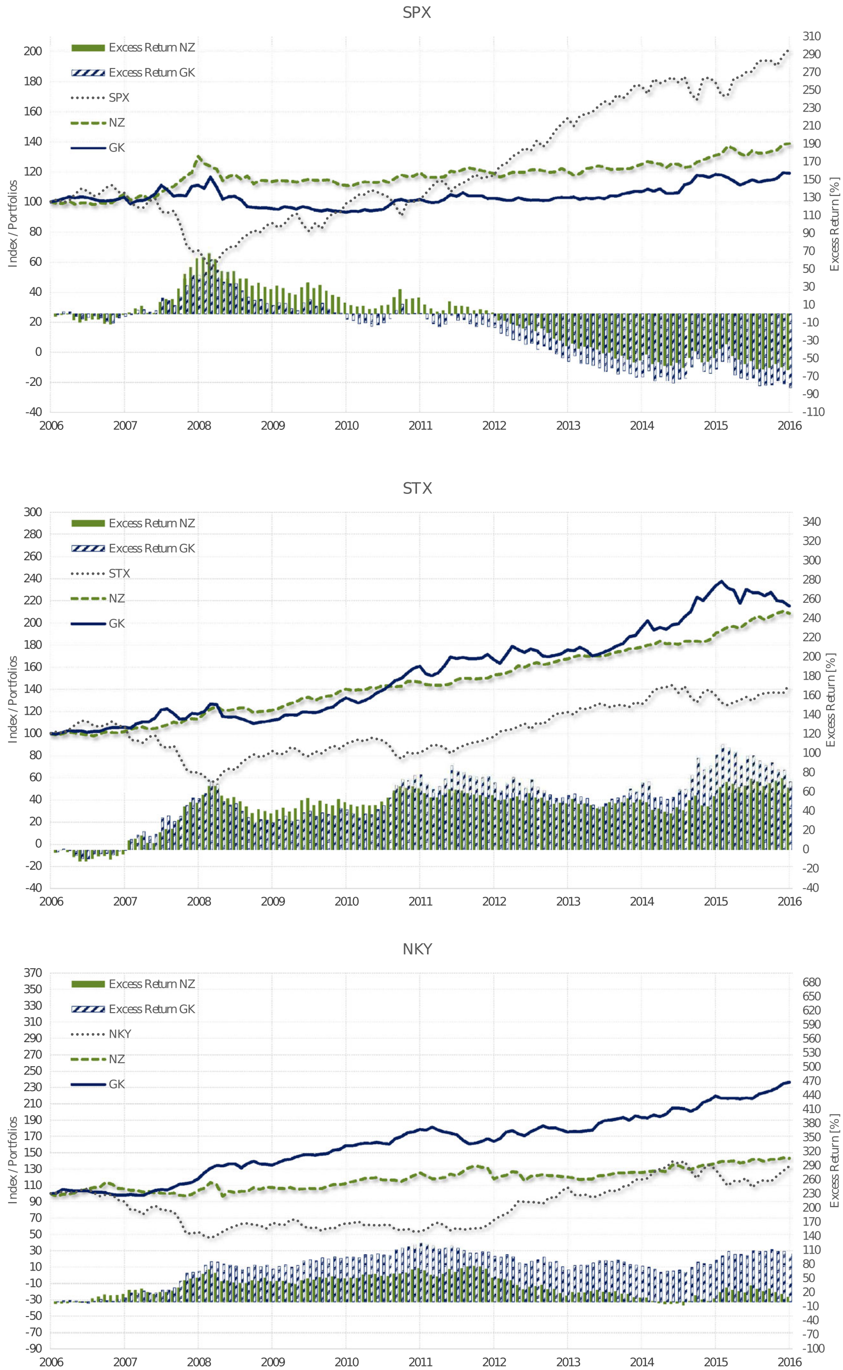

Fig. 5 Cumulative performance results of long-short multifactor portfolios. Cumulative portfolio returns appear on the left axis and excess returns in comparison to the index returns are shown as vertical bars with their corresponding values on the right axis

References

Bender, J., Wang, T.: Can the whole be more than the sum of the parts? bottom-up versus top-down multifactor portfolio construction. J. Portf. Manage. **42**(5), 39–50 (2016)

Buckle, D.: How to calculate breadth: an evolution of the fundamental law of active portfolio management. J. Asset Manage. **4**(6), 393–405 (2004)

Clarke, R., De Silva, H., Thorley, S.: Portfolio constraints and the fundamental law of active management. Financ. Anal. J. **58**(5), 48–66 (2002)

Ding, Z., Martin, RD.: Time varying IC and optimal portfolio turnover (2018)

Frazzini, A., Israel, R., Moskowitz, T.J.: Trading costs (2018)

Green, J., Hand, J.R., Zhang, X.F.: The characteristics that provide independent information about average us monthly stock returns. Rev. Financ. Stud. **30**(12), 4389–4436 (2017)

Grinold, R.C.: The fundamental law of active management. J. Portf. Manage. **15**(3), 30–37 (1989)

Grinold, R.C.: Alpha is volatility times IC times score. J. Portf. Manage. **20**(4), 9–16 (1994)

Grinold, R.C., Kahn, R.N.: Active Portfolio Management: A Quantitative Approach for Providing Superior Returns and Controlling Risk, 2nd edn. McGraw-Hill, New York (2000)

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

Lewellen, J.: The cross-section of expected stock returns. Crit. Financ. Rev. **4**, 1–44 (2015)

MacKinlay, A.C., Pástor, L.: Asset pricing models: implications for expected returns and portfolio selection. Rev. Financ. Stud. **13**(4), 883–916 (2000)

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

MSCI Global Investable Market Value and Growth Index Methodology. Methodology (2017). https://www.msci.com/eqb/methodology/meth_docs/MSCI_GIMIVGMethod_Sep2017.pdf

Pachamanova, D.A., Fabozzi, F.J.: Recent trends in equity portfolio construction analytics. J. Portf. Manage. **40**(3), 137–151 (2014)

Schäfer, J., Strimmer, K.: A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. Stat. Appl. Genet. Mol. Biol. **4**(1), 32 (2005). https://doi.org/10.2202/1544-6115.1175

Treynor, J.L., Black, F.: How to use security analysis to improve portfolio selection. J. Bus. **46**, 66–86 (1973)

Ye, J.: How variation in signal quality affects performance. Financ. Anal. J. **64**(4), 48–61 (2008)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Lars Heinrich works as a portfolio manager in the department of equity and multi-asset strategies at W&W Asset Management. He received an M.A. degree in applied science as well as a second M.A. degree in business administration with a focus on financial management at the Technical University of Freiberg. Apart from his activities as a portfolio manager, Lars Heinrich researches at the Viadrina European University of Frankfurt (Oder). His research focus lies on factor investing and alpha forecasting.

Martin Zurek works as a research assistant at the chair of finance and capital market theory of the Viadrina European University of Frankfurt (Oder). He received an M.Sc. and B.Sc. degree in business administration with a focus on finance at the Viadrina European University. Apart from his activities as a research assistant, Martin Zurek takes part in the Ph.D. program of the Viadrina European University. His research focus lies on portfolio management and asset pricing.