Quantitative Finance, 2021
Vol. 21, No. 6, 893–909, https://doi.org/10.1080/14697688.2020.1868558

Smart Alpha: active management with unstable and latent factors

C. BOUCHER†, A. JASINSKI†, P. KOUONTCHOU‡ and S. TOKPAVI*§

†EconomiX-UPL, CNRS, University Paris Nanterre, Paris, France
‡CEREFIGE, University of Lorraine, Metz, France
§Laboratoire d'Economie d'Orléans, University of Orléans, Rue de Blois, Orléans, France

(*Received 4 May 2020; accepted 18 December 2020; published online 2 February 2021*)

Factor investing has attracted increasing interest in the investment industry because purely active and passive solutions have underperformed. Its success depends critically on identifying the factors involved and timing this well, but this is hard to do because there is such a zoo of factors, and those factors and their loadings are time-varying. We thus propose an investment rule that we call 'Smart Alpha', which avoids betting on a-priori factors but focuses instead on an active approach that minimises the exposure of the portfolio to systematic sources of risk while maximising its alpha. This means our choice is to bet on alphas instead of alternative betas. We use stocks in the European STOXX 600 universe to show empirically that the Smart Alpha portfolio dominates many popular European factor investing indexes and smart beta strategies.

Keywords: Portfolio optimisation; Factor investing; Zoo of factors; Sparse latent factors; Time-varying factors

JEL Classifications: C1, C50, C61, G11, G12

1. Introduction

This article proposes an active investment approach for allocating equity portfolios. The strategy consists of betting on alphas, rather than on specified a-priori factors that the factor investing approach does. This is done through a strategy we call 'Smart Alpha', which maximises alpha, or the expected return that is uncorrelated to various systematic sources of risk, while minimising the systematic risk from exposure to these risk factors. Computationally, the alphas of stocks and their exposures to their driving factors are estimated using Sparse Principal Component Analysis (SPCA), a dimension reduction method from machine learning, coupled with a methodology for timing the optimal number of unknown relevant factors.

The efficiency of the Markowitz (1952) mean-variance optimised portfolios and of the market-capitalisation weighting derived from the Capital Asset Pricing Model (Sharpe 1964) has recently been challenged. They use a rather 'heroic' set of assumptions (stability in the expected returns and the variance-covariance matrix, no estimation risk, normal returns, no constraints on short selling, homogeneous expectations, and more) to guarantee that no other portfolios with the same risk have higher expected return. Furthermore, mean-variance optimised portfolios and market-capitalisation weighted portfolios perform relatively poorly out-of-sample (see for example Bloomfield *et al.* 1977, DeMiguel *et al.* 2009, Tu and Zhou 2011, Behr *et al.* 2012, Kourtis *et al.* 2012). The consequence is that alternative weighting schemes and heuristic approaches (equally weighted, minimum variance, most diversified portfolio, equally weighted risk contributions, risk budgeting or risk parity)† have recently been proposed and have rapidly attracted asset managers and large institutional investors under the labels 'smart beta' or 'factor investing'.‡ The idea underlying these approaches is to capture risk factors and thus risk premiums like low volatility, momentum, quality, value and size.

The rise of factor investing stems from the findings of academic research into the existence of some common risk factors beyond the market index. This strand of literature, which can be dated back to the seminal work of Fama and

*Corresponding author. Email: sessi.tokpavi@univ-orleans.fr

† See Choueifaty and Coignard (2008), Meucci (2009), Maillard *et al.* (2010).
‡ Smart beta assets under management have grown by 500 % since 2008 to $616 billion at the end of 2015 according to data from Morningstar ('Smart beta defenders dismiss fears, but doubts linger ', 11 September 2016 *Financial Times*).

French (1992), has discovered many market variables or factors that may be able to explain the cross-sectional variations of stock returns. These include the size and value factors in Fama and French (1992) and the momentum factor in Jegadeesh and Titman (1993). The result is that what is traditionally called alpha in Sharpe's single index model appears instead as beta in disguise. This raises concerns in the financial world about how relevant active portfolio management is, as the alpha of a stock measured by the average excess returns over factor risk premiums naturally shrinks as more factors are identified and are subject to bets. However, there are three arguments that qualify this assertion.

The first is that the number of factors reported in the academic literature has become so numerous that Cochrane (2011) spoke of them as a 'zoo of factors' in his 2011 presidential address to the American Finance Association.† There have been some attempts in the empirical literature to gauge the relevance of these numerous factors,‡ and Lewellen *et al.* (2010) for instance focus on biases in the cross-sectional regressions for asset pricing tests and show that the explanatory power of many of the factors suggested are spurious (see also Bai and Ng 2006, Harvey *et al.* 2016, Ahn *et al.* 2018). The second issue is that beyond this uncertainty about the number of factors and which ones really drive the cross-sectional variations of stocks, the success of factor investing is anchored in how accurately portfolio managers can time when a given factor is going to be rewarded by the market. This task is by nature complicated because the factors and loadings in markets are potentially time varying (Kelly *et al.* 2019) in response to changes in beliefs or in how markets price different stories (e.g. Shiller 2017, Gennaioli and Shleifer 2018). The third argument is that even in the ideal case when the factors rewarded are timed accurately, there is still a significant probability that stocks will have alphas that arise from stochastic mispricing relative to the asset pricing model, and from corrections of earlier over-reactions to news (de Bondt and Thaler 1985, 1987), slow adjustments to firm-specific news (Jegadeesh and Titman 1993), unanticipated increases in market illiquidity (Amihud 2002), and the state of sentiment (Lee *et al.* 1991).

From these stylised facts, we propose an approach that can be viewed as an orthogonal take to the factor investing approach, and investigate its empirical properties. We call this approach 'Smart Alpha', and it works by reducing exposures to various systematic sources of risk while maximising the expected return that is uncorrelated with these main risk factors. Rather than betting on specific a-priori factors as in the factor investing approach, we focus instead on an active approach that minimises the portfolio's exposures to systematic latent sources of risk while maximising its alpha, which is given by the expected mispriced returns from an asset pricing model. Our choice is thus to bet on alphas instead of on alternative betas as the factor investing approach does. A further reason for this choice is that empirical regularities about the predictability of alphas are reported in the literature. Zaremba *et al.* (2019) demonstrate predictive power over future returns from using alphas based on a multi-factor model with one year of data, allowing them to document an alpha momentum phenomenon for country and industry equity indexes in Europe. We hypothesise that it is likely that this phenomenon also holds for individual stocks, and our strategy intends to exploit this by maximising the portfolio's alpha.

The success of our strategy depends deeply on specifying the asset pricing model correctly in order to capture the alphas of the stocks over time as measured by the average mispriced returns. To achieve this we build on an exciting literature in statistics that focuses on estimating the number of factors in large dimensional latent factor models (Lewbel 1991, Connor and Korajczyk 1993, Donald 1997, Bai and Ng 2002, Onatski 2010, Ahn and Horenstein 2013). More precisely, we follow Bai and Ng (2002), who introduced statistics of the information criteria type into the framework of an approximate factor model to estimate the number of latent factors. Their statistics are similar to the Akaike or Bayesian information criteria used for model selection in a linear regression model, and they are a compromise between the average variance of idiosyncratic returns and the number of factors. We use their method rather than more recent ones as it is simple to understand, like any information criterion should be, and it has been shown not to be inferior to the other methods in our context, where the variance of the systematic part of stock returns is likely to be larger than the idiosyncratic part.§

The second main contribution of this paper, besides determining the optimal number of factors to help avoid specification errors in the latent factor model and guarantee that the stock's alphas are identified correctly, is to estimate the alphas of a stock based on the betas¶ extracted from the sparse principal component analysis (SPCA). SPCA is a well-known methodology in machine learning that has applications in many fields including image processing, facial recognition, gene expression data analysis, multi-scale data processing, and more. In contrast to the traditional PCA, in which each factor is a linear combination of all the input variables, this method sets some of the loadings to exactly zero, so reducing estimation error. Small loadings (or exposures of a stock to latent factors) are indeed the most subject to irrelevant variability over time because they correspond to statistical noise. Sparsity thus introduces stability into the loadings of the stocks, or its betas, and consequently into its alphas.

It should be stressed that by minimising exposures to the various systematic sources of risk, Smart Alpha has some similarities to the Betting-Against-Beta (BAB) investment rule that was popularised by the prominent work of Frazzini and Pedersen (2014). This rule states that a significant proportion of investors in the market are prohibited from using leverage or are limited in how much they may do so because of

† In the same vein William Sharpe said 'When I hear smart beta, it makes me sick' at the CFA Institute Annual conference in Seattle on 5 May 2014.

‡ Relevance means robust factors that are free of data snooping, and which approximate exposures to systematic sources of risk with consistent long-term risk premiums.

§ Empirical results, which are available from the authors upon request, show that using the more recent methods does not add value.

¶ It may be recalled that the alpha for a given stock corresponds to the average returns beyond the achieved factor risk premiums that depend on their associated loadings or betas.

margin requirements. In consequence these investors go overweight in risky or high-beta stocks instead of using leverage, and this then makes those stocks more expensive. High-beta risky stocks are consequently overvalued and are associated with low realised alpha. Obviously, a simple strategy for capitalising on this behaviour is to construct a low-beta portfolio like that used in the Smart Alpha strategy.

Our approach goes beyond the BAB investment rule in three essential points though. First, the Smart Alpha strategy not only minimises the exposure or the beta of a portfolio, but it also aims to maximise the alpha of that portfolio. It differs in this from the BAB rule, which only selects a low-beta portfolio and allows its alpha to be realised following the promise of the negative relation between betas and alphas. The second difference is that traditional implementations of the BAB strategy are based on betas that are identified as exposures to a fixed number of empirical factors (market, size, value), while our approach considers latent factors, and the number of these may vary over time as market conditions change. The third point is that a robust approach (SPCA) is used to estimate both the alphas and the betas, and this cleans out as much noise as possible.

Using stocks in the European STOXX 600 universe, we show empirically that our Smart Alpha portfolio is economically and statistically superior to many popular European factor investing indexes, the BAB portfolio, and more generally smart beta portfolios. The raw returns, average realised returns and risk-adjusted returns are, on average, higher than those of the competing portfolios, while the downside risk as measured by the maximum drawdown is lower. Moreover, we evaluate the economic value of using the SPCA to estimate the exposures and alphas of the stocks instead of the traditional PCA, and we observe significant economic gains.

The article is organised as follows. Section 2 presents the optimisation programme that underlies the Smart Alpha strategy together with a description of the latent factor model for asset pricing. Section 3 presents the method for identifying the optimal number of latent factors, and describes the SPCA methodology for estimating the exposures and alphas of the stocks. Section 4 compares the out-of-sample performances of the Smart Alpha portfolio with the MSCI factor investing indexes, BAB portfolios and popular smart beta portfolios, while Section 5 investigates the economic value of the new portfolio and its sensitivity to traditional factors. The last section concludes the paper.

2. The Smart Alpha investment strategy

This section describes our Smart Alpha investment strategy. The first part of the section describes the underlying optimisation programme, and the second part specifies the statistical asset pricing model used for estimating the main inputs, which are the alphas and betas of the stocks.

2.1. *The optimisation programme*

To describe the Smart Alpha optimisation programme, let ω be a vector of length N with elements that are the weights of a given portfolio, where N denotes the number of stocks in the capitalisation universe. The mean idea of our strategy is to bet on smart or intelligent alphas rather than on alternative betas or risk premiums. More specifically, the two objectives of the proposed strategy are to limit the exposures of the optimal portfolio to the systematic sources of risk, and to maximise its alpha given by the average mispriced returns from an asset pricing model. Starting from these two objectives, the optimisation programme can be written as follows

$$\begin{cases} \widetilde{\omega} = \arg\min\limits_{\omega} \omega' \, \Sigma_S \, \omega \\ u.c. \; \omega'\alpha \geq \epsilon, \omega \geq 0, \omega \leq \overline{\omega}, \omega' e = 1, \end{cases} \tag{1}$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)'$ is the vector of length N of the stock's alphas, $\alpha_p = \omega'\alpha$ is the portfolio's alpha, $\sigma^2_{S,p} = \omega'\Sigma_S\omega$ is the systematic portfolio's variance with Σ_S as the (N, N) systematic covariance matrix from an asset pricing model, $\epsilon > 0$ is the required portfolio's alpha, $\overline{\omega}$ is the upper-bound on the portfolio weights for the purpose of diversification, and e is the unit vector of length N. The programme incorporates an alpha targeting constraint in the low systematic risk portfolio strategy. In empirical applications, the parameter ϵ is set to an extreme upper quantile of the estimated alphas of the stocks.† We also consider stability issues by adding a turnover constraint to the programme, which is given by

$$\sum_{i=1}^{N} \left| \omega_i - \omega_i^0 \right| \leq \theta, \tag{2}$$

where ω_i^0 is the weight of stock i in the owned portfolio, and θ is the maximum value of wealth rebalancing for each optimisation, which we set to 20% in the empirical applications.

Two main remarks can be made about this optimisation programme. First, the Smart Alpha strategy limits the portfolio's exposures to the main systematic risk factors, unlike the factor investing approach, in which the philosophy is to increase exposures to empirical factors. Moreover, Smart Alpha minimises the exposures of stocks to systematic risk factors, and so is in line with the desire of investors to see more risk management practices being used in portfolio construction. Since the most recent global financial crisis there has indeed been renewed interest in de-risking investment portfolios as investors increasingly demand portfolio strategies that could protect their wealth in volatile, falling markets. Long-only equity portfolios can de-risk by reducing or targeting the exposure of the portfolio to systematic risk factors.

As the second remark, it is worth stressing that by minimising the exposure of the portfolio to the main sources of systematic risk while maximising its alpha, our strategy tries to maximise a risk-adjusted measure of performance defined as the ratio of the portfolio's alpha to its systematic volatility. It differs in this from the Markowitz view of an active portfolio strategy, in which the portfolio's expected return is maximised rather than its alpha, while the overall volatility of the portfolio is reduced rather than the systematic volatility.

† More precisely, we begin at the 99% quantile and decrease the order of the quantile by a small amount until convergence.

Interestingly, we can show that in the simplified case when the CAPM holds with the market portfolio as a single source of risk, this risk-adjusted measure is nothing else but that of Treynor (1965). In this case, and setting the risk-free rate to zero without loss of generality, we have

$$\frac{\alpha_p}{\sigma_{S,p}} = \frac{\mu_p - \beta_p \mu_M}{\sqrt{\beta_p^2 \sigma_M^2}} = \frac{\mu_p}{\beta_p \sigma_M} - \frac{\mu_M}{\sigma_M}, \tag{3}$$

where μ_p is the portfolio's expected return, $\beta_p \mu_M$ and $\beta_p^2 \sigma_M^2$ are respectively the portfolio equilibrium expected return and systematic variance from the CAPM, β_p is the portfolio beta, and μ_M and σ_M are the expected return and volatility of the market portfolio. Taking μ_M and σ_M as exogenous, it appears from (3) that maximising our ratio thus corresponds to maximising the Treynor Ratio given by μ_p/β_p.

In the following section, we specify the factor model underlying the asset pricing model that we use to estimate the inputs of our optimisation programme.

2.2. *Specification of the factor model*

The asset pricing model we consider is based on the dynamic factor model (DFM) specification for asset returns (see Geweke 1977, Sargent and Sims 1977)

$$r_t = \lambda_0 f_t + \lambda_1 f_{t-1} + \cdots + \lambda_s f_{t-s} + e_t, \tag{4}$$

where r_t is the vector of length N of returns on the stocks, f_t is a vector of length q with elements the dynamic factors, $\lambda_j, j = 0, \ldots, s$, is a matrix of dimension (N, q) with elements the exposures of stocks on the q dynamic factors, and e_t is the vector of length N of idiosyncratic or residual returns at time t. In the terminology of Bai and Ng (2007), the q dynamic factors are called primitive shocks and are supposed to drive the cross-sectional variations of the returns on the stocks.

Note that this dynamic model admits a static representation given by

$$r_t = \Lambda F_t + e_t, \tag{5}$$

where $F_t = (f_t', f_{t-1}', \ldots, f_{t-s}')'$ is a vector of length $m = q(1+s)$ of static factors, and $\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_s)$ is a matrix of dimension (N, m) of exposures to static factors. Our object of interest here is the static representation of the dynamic factor model where the m static factors are identified and estimated. To put this differently we do not focus on identifying the s primitive shocks as we do not need to do this to estimate the two elements that are important for our investment strategy, which are the alphas and betas of the stocks. With the static specification in (5), the systematic covariance matrix Σ_S in the optimisation programme (1) has the following expression

$$\Sigma_S = \Lambda \Sigma_F \Lambda', \tag{6}$$

where Λ is once again the (N, m) matrix of the exposures of the stocks to static factors and Σ_F is the (m, m) matrix corresponding to the covariance matrix of static factors. As for α, which is the vector of length N of the alphas of the stocks in the same optimisation programme, it is equal to the difference between μ, the expected value of r_t, and μ^e, the vector of expected equilibrium value, which is

$$\alpha = \mathbb{E}(e_t) = \mu - \mu^e = \mu - \Lambda \overline{F}, \tag{7}$$

where $\bar{F}$ is the vector of length m of expected returns on the m factors.

We should stress that in specification (5), we do not try to explain the cross-sectional variations of stock returns by considering the empirical factors that are identified in the literature like market, value, growth, small or large. This is because we want to avoid using spurious empirical risk factors that suffer from specification and estimation errors and that typically have a high degree of multi-collinearity. When multi-collinearity exists, there are various pitfalls that can be exacerbated when estimating betas of a stock in a multi-factor regression. These pitfalls are that the estimated regression coefficient of any one variable depends on which other predictors are included in the model; the estimated regression coefficients become less precise as more predictors are added to the model; the marginal contribution of any one predictor variable to reducing the error sum of squares depends on which other predictors are already in the model; and the hypothesis tests on coefficients may yield different conclusions depending on which predictors are in the model.

To give more insight into this last point, table 1 presents the correlations between the weekly returns of eight MSCI factor investing/smart beta European Equity indexes over the period 2001–2018. The correlations between these long-only factors are high and can be explained by the existence of a significant market component. Indeed the correlation between each of the seven non-market factors individually and the market index averages 93%. A simple solution for dealing with factor correlations is to remove the market component from the other factors. However, the issue of factor dependencies remains even if the common market component is removed. To illustrate this point, table 2 displays the correlations between the returns on the MSCI factors cleaned of the market component. This cleaning is done by extracting the market component and calculating the excess returns over the market index adjusted for the beta from each factor. We observe that while the correlations decrease on average, some of the investable factors are still highly correlated, particularly the momentum (MOM) and the minimum volatility (MV) factors, which have a correlation of 91% over our sample.† Note that this stylised fact is not specific to our sample, as we observe equally high correlations when using data over a longer period, though the factors involved are different.

Moving from the long-only MSCI factors to the long-short Fama-French factors does not solve the dependence problem between factors, as table 3 illustrates. The table reports the correlations of the weekly returns on the five Fama and French (2015) European Factors for the period from 2001 to 2018. The SMB (Small Minus Big) factor in this table is the average return on the small stock portfolios minus the average return on the big stock portfolios. The HML (High

† Statistical tests show that many of the correlations are statistically different from zero.

Table 1. Correlations of the weekly returns of factor investing/smart beta European equity indexes, 2001–2018.

Correlations	M	SC	LC	V	G	Q	MOM	MV
Market (M)	100%							
Small Capitalisations (SC)	90%	100%						
Large Capitalisations (LC)	100%	88%	100%					
Value (V)	99%	89%	98%	100%				
Growth (G)	98%	88%	98%	93%	100%			
Quality (Q)	94%	83%	94%	90%	96%	100%		
Momentum (Mom)	90%	84%	89%	85%	91%	92%	100%	
Minimum Volatility (MV)	91%	86%	91%	87%	93%	90%	98%	100%

Source: Bloomberg, Datastream, weekly data in EUR from 5 January 2001 to 25 May 2018. Computations by the authors. The European equity indexes are the MSCI Europe Total Return (M), MSCI Europe Small Capitalisation Total Return (SC), MSCI Europe Large Capitalisation Total Return (LC), MSCI Europe Value Total Return (V), MSCI Europe Growth Total Return (G), MSCI Europe Quality Total Return (Q), MSCI Europe Momentum Total Return (MOM), and the MSCI Europe Minimum Volatility Total Return (MV).

Table 2. Weekly correlations of factor investing/smart beta European equity indexes (adjusted for the market beta): 2001–2018.

Correlations	M	SC	LC	V	G	Q	MOM	MV
Market (M)	100%							
Small Capitalisations (SC)	0%	100%						
Large Capitalisations (LC)	0%	−76%	100%					
Value (V)	0%	−2%	8%	100%				
Growth (G)	0%	1%	−8%	−100%	100%			
Quality (Q)	0%	−13%	6%	−50%	51%	100%		
Momentum (Mom)	0%	18%	−23%	−40%	38%	49%	100%	
Minimum Volatility (MV)	0%	21%	−26%	−47%	45%	30%	91%	100%

Source: Bloomberg, Datastream, weekly data in EUR from 5 January 2001 to 25 May 2018. Computations by the authors. The European equity indexes are the MSCI Europe Total Return (M), MSCI Europe Small Capitalisation Total Return (SC), MSCI Europe Large Capitalisation Total Return (LC), MSCI Europe Value Total Return (V), MSCI Europe Growth Total Return (G), MSCI Europe Quality Total Return (Q), MSCI Europe Momentum Total Return (MOM), and the MSCI Europe Minimum Volatility Total Return (MV). The Small Capitalisation, Large Capitalisation, Value, Growth, Momentum, Quality and Minimum Volatility factor returns are the excess returns over the market index (adjusted for beta).

Table 3. Correlations of weekly returns of the five Fama-French European factors, 2001–2018.

Correlations	Market	SMB	HML	RMW	CMA
Market	100%				
SMB	−66%	100%			
HML	13%	−6%	100%		
RMW	−10%	3%	−50%	100%	
CMA	−30%	16%	38%	−18%	100%

Source: Fama-French European five-factor model, weekly data from 5 January 2001 to 25 May 2018. Computations by the authors.

Minus Low) factor is the average return on the value portfolios minus the average return on the growth portfolios. The RMW (Robust Minus Weak) factor is the average return on the robust operating profitability portfolios minus the average return on the weak operating profitability portfolios. The CMA (Conservative Minus Aggressive) factor is the average return on the conservative investment portfolios minus the average return on the aggressive investment portfolios. The results show that the correlations between the market and the other Fama-French factors, and between the individual Fama-French factors are relatively low. Nevertheless, these factors are not completely orthogonal, since some correlations appear high in absolute terms and are statistically significant, especially the correlation of −66% between the market and SMB factors, and that of −50% between the RMW and HML factors.

Since latent factors extracted using statistical methods for dimension reduction are by construction orthogonal and do not suffer from this drawback, we consider the m factors in (5) as unobservable, and subsequently propose a methodology for estimating these factors along with our optimisation inputs Σ_S in (6) and α in (7). We begin by presenting the method for estimating the optimal number m of factors in our factor representation. This issue is crucial as accurate estimations of both inputs require a correct specification of the factor model, which depends heavily on m.

3. Estimating the latent factor model by sparse principal component analysis

This section presents the methodology used to estimate the latent factor model in (5). The first part of the section looks at estimating the optimal number m of latent factors that drive the cross-sectional variations of the returns on stocks. Taking this estimate, the second part describes the Sparse Principal Component Analysis (SPCA), the machine learning methodology we estimate the latent factor model with, which provides the main inputs for the optimisation programme in (1).

3.1. Estimating the optimal number of latent factors

The literature has discussed widely how to infer the number of significant factors in approximate latent factor models (Stock and Watson 1999, Connor and Korajczyk 1993, Bai and Ng 2002, etc.). Bai and Ng (2002) developed statistics of the information criteria type to estimate the number of latent factors. Their statistics are a compromise between the average variance of idiosyncratic returns and the number of factors, and so they share some similarities with the Akaike or Bayesian information criteria (AIC, BIC) that are usually mobilised for model selection in regression models. More recently, Ahn and Horenstein (2013) introduced two eigenvalue ratio statistics that attain their maximal values at an unknown number of latent factors. The rationale of these statistics arises because the leading eigenvalues of the covariance matrix grow without bounds as the number of stocks increases, while the remaining values are bounded.†

As already stressed, we use the approximate latent factor model of Bai and Ng (2002) as it is simple to understand and is not inferior to the other methods in our context of a factor model for stock returns. We may rewrite the factor model from (5) as follows

$$R = F\Lambda' + E, \tag{8}$$

where R is the (T, N) matrix of returns, F is the (T, m) matrix of returns to latent factors, Λ is the (N, m) matrix of loadings, and E is the matrix of residual returns with the dimension (T, N).

We denote as V the (m, m) diagonal matrix with the m largest eigenvalues of RR' as elements. $\widetilde{F}$ is the principal component estimate of F under the normalisation $T^{-1}F'F = I_m$, with I_m as the identity matrix of dimension m. Then the elements of $\widetilde{F}$ are those of the eigenvector matrix associated with the m largest eigenvalues of the matrix RR', multiplied by $\sqrt{T}$. The estimate of the factor loadings matrix Λ is $\widetilde{\Lambda} = R'\widetilde{F}/T$. The estimated matrix (T, N) of residual returns $\widetilde{E}$ with the elements $\widetilde{e}_{it}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$ thus corresponds to

$$\widetilde{E} = R - \widetilde{F}\widetilde{\Lambda}'. \tag{9}$$

Under this framework, Bai and Ng (2002) propose that m be estimated using statistics like information criteria. The idea is to find a balance between residual variance, which diminishes mechanically with the number of factors, and the complexity of the model that is increasing in the number of factors. We favour the following version of these statistics‡

$$IC\,(k) = \ln(V(k, \widetilde{F}^{(k)})) + k\left(\frac{N+T}{NT}\right)\ln(C_{TN}^2), \tag{10}$$

where $V(k, \widetilde{F}^{(k)})$ is the average residual variance across stocks and time when the number of factors is set to k,

$$V(k, \widetilde{F}^{(k)}) = \frac{1}{TN}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{e}_{it}^2, \tag{11}$$

and $C_{TN} = \min\{\sqrt{N}, \sqrt{T}\}$. The estimated value $\widehat{m}$ of the number of latent factors m corresponds to the value of k that minimises the information criterion $IC(k)$, so

$$\widehat{m} = \underset{k \le k_{\max}}{\arg\min}\, IC\,(k)\,, \tag{12}$$

with $k_{\max}$ as the maximum number of factors, which we set to 50 in the empirical section. Under some regularity conditions, Bai and Ng (2002) show that $\Pr(\widehat{m} \to m) \to 1$, as $T, N \to \infty$.

An illustration of this estimation procedure is given below, using the daily returns of all the stocks in the STOXX 600 universe over the period from November 2001 to May 2018. At the end of each month, we use the most recent one-year data (260 observations) for all the stocks in the capitalisation universe and run the methodology described above to estimate the number of latent factors $\widehat{m}$. This process is iterated by repeatedly moving the estimation window one month forward, taking in the data for a new month and dropping the data for the earliest month, until the last observation is reached.

Figure 1 gives the time series dynamics of the number of estimated latent factors throughout the rolling-window procedure, with crisis periods highlighted. The crisis periods are the dot-com bubble financial crisis from September 2000 to March 2003, the global financial crisis from June 2007 to March 2009, the European sovereign debt crisis from April 2011 to June 2012, and the fear of a hard landing for China from June 2015 to June 2016. The average number of estimated latent factors over the whole sample is four. We observe a substantial variation over time, with a clear-cut shape that corresponds to a significant increase at the beginning of crisis periods, and a decrease at the exit from them. The most important sharp increase occurred during the global financial crisis, when the number of latent factors extracted jumped from three to six significant factors in the two months from 4 September 2008 to 3 November 2008. This dynamic corroborates the results in Calomiris *et al.* (2012), who studied the sensitivities of stock returns to factors in crisis periods. They isolate three factors that explain the cross-sectional variation of stock returns during the 2007–2008 global financial crisis, looking beyond the usual factors that are associated with stock returns. These factors, which they called ‘crisis shocks’, measure three market or economic states, which are a collapse of global trade, a contraction of the credit supply, and downwards pressure on the equity of firms because of selling. Most importantly, they show that these factors are not in play in various placebo samples that correspond to non-crisis periods. This then implies that the number of factors increases in crisis periods.

It is worth stressing that this increase in the number of factors in a crisis period is also observed for the other information criteria in Bai and Ng (2002), and in the method for estimating the number of factors in Ahn and Horenstein (2013). At first

† For other references in this literature, see Lewbel (1991), Connor and Korajczyk (1993), Donald (1997), and Onatski (2010).

‡ The results available from the authors upon request show that a proposed investment strategy based on the other information criteria in Bai and Ng (2002) are less economically valuable. This is because they lead to optimal portfolios with lower realised alphas on an annual basis.

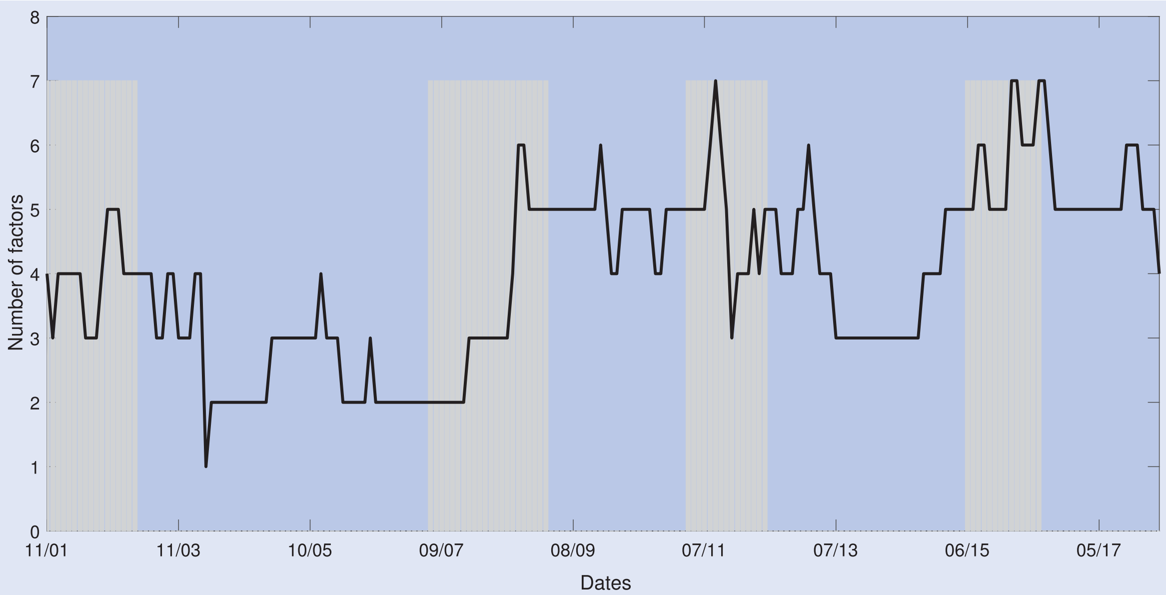

Figure 1. Dynamics of the estimated number of latent factors. Source: Bloomberg, daily data from 30 November 2001 to 28 May 2018. Computations by the authors. The four regions highlighted correspond respectively to the dot-com crisis period from 29 September 2000 to 31 March 2003, the global financial crisis period from 29 June 2007 to 31 March 2009, the European sovereign debt crisis from 29 April 2011 to 29 June 2012, and the fear of a hard landing for China from June 2015 to June 2016.

glance this result might seem to contradict the empirical regularity that correlations increase in times of crisis. However, this regularity stems rather from the increase in the explanatory power of the first factor, notably the market, relative to the other factors in times of crisis, and is not incompatible with an increase in the number of factors in these periods. Figure 2 represents the relative explanatory power of the significant alternative factors, as measured by the ratio of the sum of the variances or eigenvalues of these factors over the variance of the first, market, factor. This figure shows that this power decreases in crisis periods, leading to the conclusion that the relative part of the variance of the market factor increases significantly in these periods.

3.2. Description of the sparse principal component analysis

The success of the investment strategy that we develop in this paper is based on estimating the alphas and betas of stocks efficiently. This section describes the Sparse Principal Component Analysis (SPCA) methodology that we use to estimate this set of inputs for our optimisation programme. Historically, this method aims to alleviate one of the main pitfalls of the traditional PCA, which is that the factors are hard to interpret. In PCA, the loadings or exposures to factors are non-zero, with the consequence that each factor is a linear combination of all the input variables. This makes it harder to interpret the factors extracted.

To achieve interpretability beyond data representation, SPCA sets some of the loadings, and hence the number of input variables that contribute to the variance of each factor, to zero. This can be done in an ad hoc way by choosing a threshold value and zeroing all the small loadings with absolute values lower than the threshold, but this approach has been criticised in the literature (Cadima and Jolliffe 1995). More efficient methods (Jolliffe *et al.* 2003, Zou *et al.* 2006, etc.) have been proposed for introducing sparsity to the vector or matrix of loadings through penalisation methods such as the Lasso (least absolute shrinkage and selection operator) of Tibshirani (1996) or the Elastic Net of Zou and Hastie (2005). These introduce a bound for the sum of the absolute values of the loadings on each factor, forcing some of them to become zero.†

In our context of estimating a latent factor model, SPCA has more virtues than just interpretability. First, by setting some of the loadings to exactly zero, it helps reduce estimation error. Indeed it is small loadings, or small exposures of stocks to latent factors, that are the most subject to irrelevant variability over time, because they correspond to statistical noise. Sparsity thus introduces stability in the loadings of a stock, which are its betas, and consequently in its alpha. Second, it makes sense to hypothesise that not all stocks are exposed to all factors, and therefore that only a certain set of stocks contributes to explaining the variability of each given factor. So if one of the latent factors extracted has a high level of correlation with, say, the small capitalisation factor as an empirical factor, it is reasonable to expect that stocks with large capitalisation will not load on this factor.

Here we use the SPCA methodology of Wu and Chen (2016), which is well suited for large dimensional problems. This method is introduced by the authors as an alternative to the SPCA approach in Zou *et al.* (2006). To describe the approach, we may consider once again our latent

† There are many variants of the Sparse PCA in the literature. See for example the SPCA method of Zou *et al.* (2006), the direct SPCA approach in d'Aspremont *et al.* (2007), the branch and bound method in Farcomeni (2009), the GPower method of Journée *et al.* (2010), the iterative elimination algorithm in Wang and Wu (2012), or the thresholding approach in Wu and Chen (2016). For a review, see Feng *et al.* (2016).

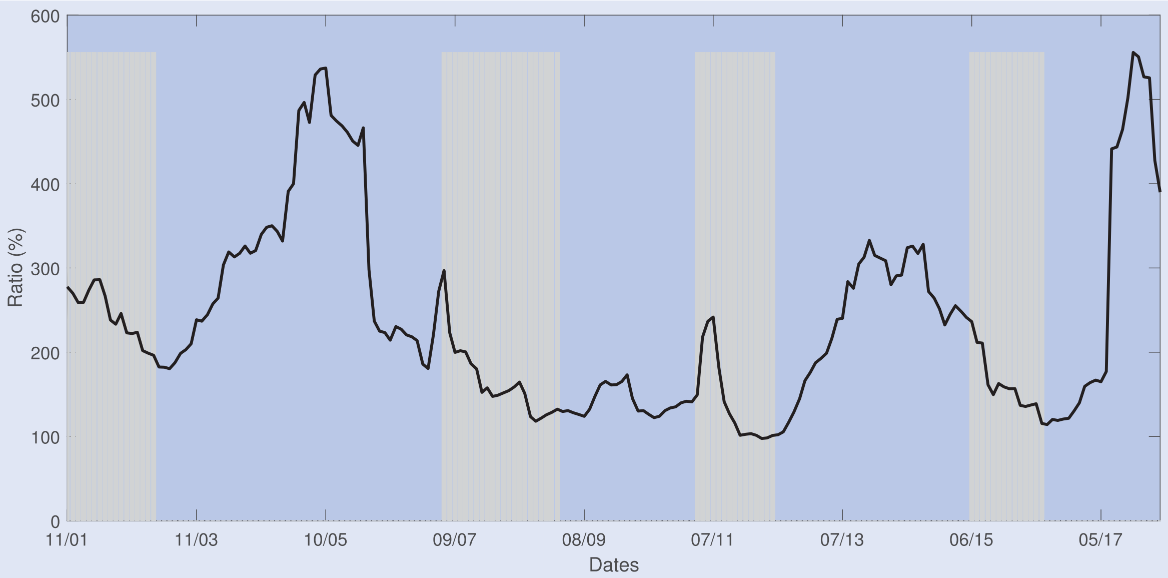

Figure 2. Relative explanatory power of significant alternative factors. Source: Bloomberg, daily data from 30 November 2001 to 28 May 2018. Computations by the authors. The relative explanatory power of the significant alternative factors is computed as the ratio of the sum of the variances of these factors over the variance of the first, market, factor. The four regions highlighted correspond respectively to the dot-com crisis period from 29 September 2000 to 31 March 2003, the global financial crisis period from 29 June 2007 to 31 March 2009, the European sovereign debt crisis from 29 April 2011 to 29 June 2012, and the fear of a hard landing for China from June 2015 to June 2016.

factor model given by

$$R = F\Lambda' + E, \tag{13}$$

where R is again the matrix (T, N) of returns on stocks, F is the matrix $(T, \widehat{m})$ with columns given by the T returns of each latent factor, while $\widehat{m}$ is the number of optimal factors obtained using the information criterion (see equations (10)–(12)). The matrix Λ is of dimension $(N, \widehat{m})$ with the rows containing the betas or exposures of each of the N stocks to the $\widehat{m}$ factors, and E is the matrix (T, N) of idiosyncratic or residual returns.

Given that this specification is based on the optimal number of latent factors, the traditional PCA methodology can be used to get the estimate $\widehat{\Lambda}_{pca}$ from the singular value decomposition of the matrix of returns R, with $\widehat{\Lambda}_{pca}$ corresponding to the first $\widehat{m}$ columns of the matrix V, with

$$R = U\Omega V', \tag{14}$$

where the orthogonal matrices U and V are of dimensions (T, N) and (N, N) respectively, and Ω is a diagonal matrix of dimension (N, N). As shown by Zou *et al.* (2006), the traditional PCA estimate of Λ as described above can be recast in a linear ridge regression framework. To do this, let $A = [a_1, \ldots, a_{\widehat{m}}]$ and $B = [b_1, \ldots, b_{\widehat{m}}]$ be the matrices of dimension $(N, \widehat{m})$. For any $\psi > 0$, consider the following constrained ridge regression

$$(\widehat{A}, \widehat{B}) = \underset{(A,B)}{\arg\min} \sum_{j=1}^{\widehat{m}} \left\| Ra_j - Rb_j \right\|_2^2 + \psi \left\| b_j \right\|_2^2, \text{ u.c.} A'A = I_{\widehat{m}}, \tag{15}$$

where R is the matrix (T, N) of the stock's returns, $\psi > 0$ is the ridge parameter, and $I_{\widehat{m}}$ is the identity matrix of dimension $\widehat{m}$. As shown by Zou *et al.* (2006), the estimated matrix $\widehat{B} = [\widehat{b}_1, \ldots, \widehat{b}_{\widehat{m}}]$ is equal up to some normalisations to the estimated matrix $\widehat{\Lambda}_{pca} = [\widehat{\lambda}_{1,pca}, \ldots, \widehat{\lambda}_{\widehat{m},pca}]$ from the PCA, so

$$\widehat{\lambda}_{j,pca} = \frac{\widehat{b}_j}{\left\| \widehat{b}_j \right\|_2}, \quad j = 1, \ldots, \widehat{m}. \tag{16}$$

Following from this result, Wu and Chen (2016) add a thresholding constraint on this programme to obtain the sparse loadings of the stock, yielding

$$\begin{cases} (\widetilde{A}, \widetilde{B}) = \underset{(A,B)}{\arg\min} \sum_{j=1}^{\widehat{m}} \|Ra_j - Rb_j\|_2^2 + \psi \|b_j\|_2^2 \\ \text{u.c. } A'A = I_{\widehat{m}} \; B(i,j) = 0 \; if \; G(i,j) = 0, \end{cases}, \tag{17}$$

where $G(i,j)$ is a sparse regularisation matrix with dimensions $(N, \widehat{m})$. The choice of the matrix G is crucial in this framework as it conditions the level of sparsity for the estimated loadings. Here we follow the suggestion of estimating G from a decision matrix based on the matrix of correlation of the stock's returns. The underlying idea is that a stock with a large variance should load on a significant factor, and the other stocks that are highly correlated with this benchmark stock should also load on this factor, and those that are less correlated should not.†

An equivalent form of the constrained penalised regression (17) is given by the authors. By letting D_j be the diagonal matrix with $D_j(i,i) = G(i,j)$, they show that (17) does indeed correspond to a constrained ridge regression

$$(\widetilde{A}, \widetilde{B}) = \underset{(A,B)}{\arg\min} \sum_{j=1}^{\widehat{m}} \left\| Ra_j - RD_jb_j \right\|_2^2 + \psi \left\| b_j \right\|_2^2$$
$$\text{u.c. } A'A = I_{\widehat{m}}. \tag{18}$$

† See the reference for more details on this issue.

with the solutions

$$\widetilde{b}_j = \left(D_j R' R D_j + \psi I_N\right)^{-1} D_j R' R \widetilde{a}_j. \tag{19}$$

The connection to the hard-thresholding rule arises in a large dimensional setting by letting $\psi \to \infty$, with (19) becoming

$$\widetilde{b}_j = D_j R' R \widetilde{a}_j, \tag{20}$$

and so the sparse loadings are defined as

$$\widetilde{\lambda}_{j,spca} = \frac{\widetilde{b}_j}{\|\widetilde{b}_j\|_2}, \quad j = 1, \ldots, \widehat{m}. \tag{21}$$

Because in (20) $\widetilde{b}_j$ depends on $\widetilde{a}_j$, Wu and Chen (2016) proposes an iterative algorithm that provides solutions to (17) or equivalently (18) by solving for B with A kept fixed, or for A with B fixed.† The full SPCA algorithm of Wu and Chen (2016) is displayed in Appendix A.

With the estimated matrix $(N, \widehat{m})$ of sparse loadings $\widetilde{\Lambda}_{spca} = (\widetilde{\lambda}_{1,spca}, \ldots, \widetilde{\lambda}_{\widehat{m},spca})$, the matrix $(T, \widehat{m})$ of the estimated factors is given by

$$\widetilde{F}_{spca} = R \widetilde{\Lambda}_{spca}. \tag{22}$$

An empirical view of the estimated factors obtained with the SPCA methodology and their link with a-priori factors in the academic literature is given in Table B1 in Appendix B, which presents the a-priori factors for each year that are most correlated with the estimated latent factors (best correlated factor with the fitted latent factor for each year). The estimated latent factors are presented in each column and the absolute value of the correlation coefficient is presented below the name of the a-priori factor. We considered a list of 49 potential a-priori factors available at a daily frequency, from oil price returns and equity sector returns to credit, term and sovereign spreads, implied volatilities, gold price returns and returns on typical indexes. The first point to note is that the main latent factor appears each year to be perfectly correlated with the market portfolio, which is the MSCI Europe index. The second is that we observe an instability in identifying the other latent factors since the most correlated a-priori factors change across the years. Moreover, the other latent factors present weak correlations with a-priori factors in most of the cases. These results explain why factor identification or timing is very difficult, and they justify our approach of betting on alpha rather than on these alternative risk premiums that were highlighted in the empirical literature on asset pricing.

Note that the degree of sparsity of each factor estimated is displayed in Table B2 in Appendix B. For each year and factor, it corresponds to the proportion of the loadings of stocks that have estimated betas or exposures equal to zero. We observe that the degree of sparsity evolves over time and there is a lot of variation across factors. Interestingly, the degree of sparsity of the first factor is always equal to zero, though this result is expected as the first identified factor is nothing but the market. Overall, the degree of sparsity for each year decreases with the explanatory power of the factor. This means that the less a factor explains the variability of the data, the fewer stocks are exposed to it.

Lastly, let us note that with $\widetilde{\Lambda}_{spca} = (\widetilde{\lambda}_{1,spca}, \ldots, \widetilde{\lambda}_{\widehat{m},spca})$, the estimate of the vector $\alpha = (\alpha_1, \ldots, \alpha_N)'$ of length N of stock's alphas is given by

$$\widetilde{\alpha}_{spca} = \bar{\widetilde{E}} = \bar{R} - \bar{\widetilde{F}}_{spca} \widetilde{\Lambda}'_{spca}, \tag{23}$$

where $\bar{\widetilde{E}}$, $\bar{R}$ and $\bar{\widetilde{F}}_{spca}$ are the sample means of $\widetilde{E}$, R and $\widetilde{F}_{spca}$, respectively. The feasible version of our Smart Alpha optimisation programme in (1) becomes

$$\begin{cases} \widetilde{\omega} = \arg\min_{\omega} \omega' \widetilde{\Sigma}_{S,spca} \omega \\ u.c.\ \omega' \widetilde{\alpha}_{spca} \geq \epsilon, \omega \geq 0, \omega \leq \overline{\omega}, \omega' e = 1, \end{cases} \tag{24}$$

where $\widetilde{\Sigma}_{S,spca}$ is the systematic covariance matrix given by $\widetilde{\Sigma}_{S,spca} = \widetilde{\Lambda}_{spca} \widetilde{\Sigma}_{F,spca} \widetilde{\Lambda}'_{spca}$, with $\widetilde{\Sigma}_{F,spca}$ as the covariance matrix of the $\widehat{m}$ factors.

Note that for our optimisation programme in (24), the objective function is convex, which ensures that there is a unique global solution as soon as the systematic covariance matrix $\widetilde{\Sigma}_{S,spca}$ is positive-definite. Moreover, even in the opposite case, the use of the constraints on the weights makes it possible to obtain a global solution. Indeed Jagannathan and Ma (2003) demonstrated that using the positivity constraint on the weights is equivalent to regularising the covariance matrix $\widetilde{\Sigma}_{S,spca}$ in the quadratic form of the objective function. In practice, for the first optimisation of our backtesting exercise, we use the equally weighted portfolio as the initial solution, and for the following optimisations we use the portfolio held at the time of the optimisation.

4. The Smart Alpha and competing portfolios: comparison of performance

This section evaluates the Smart Alpha portfolio strategy empirically looking at the European equity market. It compares the performance of the Smart Alpha portfolio with those of European factor investing indexes and some competing smart beta portfolios. We start by describing the database and the methodology used for the evaluation, and then we analyse the performance profiles of the strategies.

We consider the European stock market here because it has valuable characteristics in its size, liquidity, diversity of market conditions, and representativeness. Moreover it offers a significant regional diversification through different stock exchanges that operate independently, as 17 countries are represented within the STOXX 600 index, and so also through multiple currencies and then central banks that may drive the risks and returns of stocks. That means this market provides fertile ground for active strategies that try to capitalise on the alphas of stocks. The STOXX 600 serves as a benchmark for numerous actively managed funds and it underlies a variety of financial products.

† Note that for B fixed, A is obtained via the reduced rank Procrustes rotation. Formally, $A = \widetilde{U}\widetilde{V}'$ with $\widetilde{U}$ and $\widetilde{V}$ as the orthogonal matrices from the singular value decomposition of $R'RB$, so $R'RB = \widetilde{U}\widetilde{\Omega}\widetilde{V}'$. For more discussion, see Zou *et al.* (2006).

We could equally well have used the US stock market for our backtesting exercise but we chose not to because over recent years, the US stock market has been dominated by just a few individual stocks from the technology industries. Stocks of the five biggest companies (Alphabet, Apple, Facebook, Amazon and Microsoft), known collectively as the GAFAM stocks, have been one of the main drivers of the entire US stock market for almost a decade. At 24.0%, their index weight represents nearly a quarter of the S&P 500 index as at 31 August 2020. Under these conditions, the results of an active stock-picking strategy will be very sensitive to the over or under-weighting of these few stocks.

Last but not least, the European market has experienced more episodes of financial disturbance than the United States market, with the euro area debt crisis from 2010 to end of 2012 and the Brexit crisis in 2016, which makes the backtesting exercises more robust.

4.1. Data and the evaluation methodology

The database comes from Bloomberg and Datastream, and includes the daily returns of all the constituents of the STOXX 600 index from 4 December 2000 to 28 May 2018. The database contains $N = 1278$ stocks, which is all the stocks in the composition of the STOXX 600 index since 4 December 2000, and $T = 4561$ daily observations. This database therefore takes account of survivorship bias, as it is not limited to only those stocks that are in the capitalisation universe at the end of the sample.

We use a rolling-window approach to this dataset to generate out-of-sample returns for the Smart Alpha portfolio. More precisely, we use the $n = 260$ daily returns that precede the date we set for the first optimisation of 30 November 2001 to find the solution $\widetilde{\omega}$ for the optimisation programme in (24), restricting the investment universe to only the 600 stocks in the capitalisation universe at the optimisation date, and setting the upper-bound $\overline{\omega}$ to the value 2%, meaning there are around 50 active stocks. This solution is implemented with a delay of one day, and the ex-post or realised daily returns for the strategy are recorded from the subsequent days with a holding period of one month.

This process is iterated by repeatedly moving the estimation window forward one month by including the data for a new month and dropping the data for the earliest month, until the last observation is reached. Traditional performance measures are computed for the realised returns net of transaction costs, with the proportional transaction cost set to 25 basis points per transaction.

4.2. The Smart Alpha portfolio: comparison with factor investing indexes and the BAB benchmark

Since our Smart Alpha portfolio strategy is by design orthogonal to factor investing, we compare its performances to those of eight MSCI factor investing indexes, which are the MSCI Europe Small Capitalisation Total Return (SC), the MSCI Europe Large Capitalisation Total Return (LC), the MSCI Europe Value Total Return (V), the MSCI Europe Growth Total Return (G), the MSCI Europe Quality Total Return (Q), the MSCI Europe Momentum Total Return (MOM), the MSCI Europe Minimum Volatility Total Return (MV), and the MSCI Europe Index (E). We also include the BAB strategy in the comparison as it shares some similarities with our strategy, as discussed in the introduction. The related optimal portfolio is obtained by minimising the exposures to the three Fama-French factors of market, size, and value.

The realised or ex-post performances of the Smart Alpha portfolio, the eight factor investing indexes, and the BAB portfolio are displayed in table 4. For each strategy, the table displays absolute performance measures given by the raw return, the annualised average return, the annualised volatility, the Sharpe ratio and the maximum drawdown. Performance measures relative to the STOXX 600 index and given by beta, annualised residual risk, annualised alpha, annualised average excess returns and the Appraisal ratio are also displayed.

Observing the absolute performance measures except volatility and maximum drawdown shows the best-performing factor investing index to be the MSCI Europe Small Capitalisation Total Return (SC). This index takes the highest values for the raw return at 410.99%, the annualised average return at 10.36%, and the Sharpe ratio at 0.61. However, this is beaten by our Smart Alpha portfolio (SA-SPCA), which delivers higher values for these three evaluation criteria, producing raw return of 584.01%, annualised average return of 12.33%, and a Sharpe ratio of 1.16. The Smart Alpha also delivers the best performance for the risk measures given by volatility and maximum drawdown. It has annualised volatility of 10.6% and maximum drawdown of −38.21%, which are lower than the 14.39% for annualised volatility recorded by the MSCI Europe Minimum Volatility Total Return and the maximum drawdown of −39.75% of the MSCI Europe Quality Total Return. The comparison with the BAB portfolio shows that our strategy performs better in these absolute metrics, except for volatility, which is of same order for both strategies.

This outperformance seems consistent when the measures of relative performance are considered. The annualised average excess return over the STOXX 600 index is 11.09%, which is much higher than the values for the competing strategies. Correcting the annualised average excess returns for the exposure to the market, or beta, does not eliminate this outperformance as the annualised alpha is instead 9.74%, which is still high. It may be noted that the beta of the Smart Alpha portfolio has the lowest value, which may partly explain its resilience in crises, as shown by it having the lowest value for maximum drawdown. This is also the case for the BAB portfolio. Finally, the value of the Appraisal ratio is 1.58 for the Smart Alpha portfolio, which is the highest value of any of the strategies.

4.3. Smart Alpha portfolio: comparison with popular smart beta portfolios

In this section, we compare the Smart Alpha portfolio with three smart beta portfolios and the market portfolio (STOXX 600). The first smart beta portfolio is the popular minimum

Table 4. Ex-post performances of the Smart Alpha portfolio, MSCI factor investing indexes and BAB portfolio.

	SA-SPCA	SC	LC	V	G	Q	MOM	MV	E	BAB
Raw Return (%)	584.01	410.99	106.41	104.54	137.67	249.22	351.04	196.74	123.21	347.18
Annualised Average Return (%)	12.33	10.36	4.48	4.42	5.37	7.85	9.53	6.80	4.98	9.47
Annualised Volatility (%)	10.60	16.89	19.77	21.41	18.04	16.80	17.58	14.39	19.38	10.56
Sharpe Ratio	1.16	0.61	0.23	0.21	0.30	0.47	0.54	0.47	0.25	0.89
Maximum Drawdown (%)	−38.21	−65.44	−57.26	−64.78	−51.20	−39.75	−50.84	−50.50	−58.22	−43.01
Beta	0.45	0.78	1.02	1.09	0.91	0.82	0.82	0.72	1.00	0.44
Annualised Residual Risk (%)	6.14	7.47	1.58	3.62	3.45	5.49	7.64	3.76	0.75	6.32
Annualised Alpha (%)	9.74	6.41	−0.28	−0.47	0.92	3.71	5.49	2.99	0.23	7.01
Annualised Average Excess Return (%)	11.09	8.68	−0.53	−0.65	1.25	5.29	7.61	3.68	0.49	7.53
Appraisal Ratio	1.58	0.86	−0.17	−0.13	0.27	0.67	0.71	0.79	0.31	1.11

Source: Bloomberg, daily data from 3 December 2001 to 28 May 2018. Computations by the authors. SA-SPCA refers to our Smart Alpha portfolio strategy based on the SPCA and the optimal number of factors, SC the MSCI Europe Small Capitalisation Total Return, LC the MSCI Europe Large Capitalisation Total Return, V the MSCI Europe Value Total Return, G the MSCI Europe Growth Total Return, Q the MSCI Europe Quality Total Return, MOM, the MSCI Europe Momentum Total Return, MV the MSCI Europe Minimum Volatility Total Return, and E the MSCI Europe Index. BAB is the Betting-Against-Beta portfolio strategy. The relative statistics (beta, residual risk, alpha, average excess returns, Appraisal Ratio) are computed with the STOXX 600 as benchmark.

Table 5. Ex-post performances of the Smart Alpha and popular Smart Beta strategies.

	SA-SPCA	RP	EW	MV-OPT	Benchmark
Raw Return (%)	584.01	257.90	226.40	368.90	114.79
Annualised Average Return (%)	12.33	8.01	7.41	9.79	4.73
Annualised Volatility (%)	10.60	17.39	19.07	9.94	19.35
Sharpe Ratio	1.16	0.46	0.39	0.98	0.24
Maximum Drawdown (%)	−38.21	−60.95	−63.87	−41.15	−58.69
Beta	0.45	0.87	0.96	0.42	
Annualised Residual Risk (%)	6.14	4.15	4.63	5.61	
Annualised Alpha (%)	9.74	3.62	2.80	7.34	
Annualised Average Excess Return (%)	11.09	5.52	4.64	7.94	
Appraisal Ratio	1.58	0.87	0.60	1.31	

Source: Bloomberg, daily data from 3 December 2001 to 28 May 2018. Computations by the authors. SA-SPCA refers to our Smart Alpha portfolio strategy based on the Sparse PCA and the optimal number of factors, RP the risk parity portfolio, EW the equally weighted portfolio, MV-OPT the optimised minimum volatility portfolio, and Benchmark the STOXX 600 Index.

volatility portfolio, which is obtained as

$$\begin{cases} \widetilde{\omega} = \underset{\omega}{\arg\min}\, \omega' \,\widetilde{\Sigma}\omega \\ u.c.\ \omega \geq 0, \omega \leq \overline{\omega},\ \omega' e = 1, \end{cases} \tag{25}$$

with $\widetilde{\Sigma}$ as an estimate of the covariance matrix of stock returns. We do not consider the sample covariance matrix, which is not well conditioned in our case, because the number of stocks ($N = 600$) is higher than the length of the sample ($n = 260$). We instead use the shrinkage covariance matrix estimate in Ledoit and Wolf (2003). The goal of this is to exploit the usual bias-variance trade-off by reducing the variance or instability of the sample covariance matrix, at the cost of a small bias.

The second smart beta alternative is the risk parity portfolio. This smart beta portfolio allocates wealth on a risk-weighted basis to avoid the pitfalls of the traditional capitalisation weighting scheme. The ultimate goal is to equalise each stock's total contribution to the overall risk in order to increase portfolio diversification. We consider here the simple but effective risk parity portfolio, with the weight $\widetilde{\omega}_i$ on each stock i equal to

$$\widetilde{\omega}_i = \frac{\frac{1}{\widetilde{\sigma}_i}}{\sum_{i=1}^{N} \frac{1}{\widetilde{\sigma}_i}}, \tag{26}$$

where $\widetilde{\sigma}_i$ is the sample standard deviation for the returns of stock i. Lastly, we consider the equally weighted portfolio as a third alternative.

Table 5 displays the performance metrics of the Smart Alpha portfolio, the three competing smart beta portfolios, and the STOXX 600 index as the benchmark. Overall, the Smart Alpha portfolio has the best performances for raw return, annualised average return, annualised average excess return, annualised alpha and Appraisal ratio. Looking at the risk metrics displayed, we observe that the annualised volatility of the Smart Alpha portfolio is higher than that of the optimised minimum volatility portfolio (MV-OPT). This result is to be expected and naturally arises because our portfolio only minimises the systematic part of the volatility, while the minimum volatility portfolio minimises the total volatility, both systematic and specific. It can also be linked to the alpha constraint underlying our strategy that tightens the region in which the minimum systematic risk portfolio is sought. It should be noted that the Smart Alpha portfolio is better for

Table 6. SPCA versus PCA: performance comparison.

	SA-SPCA	SA-PCA
Raw Return (%)	584.01	466.36
Annualised Average Return (%)	12.33	11.05
Annualised Volatility (%)	10.60	10.63
Sharpe Ratio	1.16	1.04
Monthly Turnover (%)	20.56	22.15
Maximum Drawdown (%)	−38.21	−38.83
Beta	0.45	0.46
Annualised Residual Risk (%)	6.14	5.90
Annualised Alpha (%)	9.74	8.42
Annualised Average Excess Return (%)	11.09	9.54
Appraisal Ratio	1.58	1.42

Source: Bloomberg, daily data from 3 December 2001 to 28 May 2018. Computations by the authors. SA-SPCA refers to our Smart Alpha portfolio strategy based on the Sparse PCA and the optimal number of factors, SA-PCA is its analogue based on the traditional non-sparse PCA.

maximum drawdown than the minimum variance portfolio. The explanation for this result is that alpha is about return and dampens the severity of losses. Lastly, our portfolio using an active strategy has the highest value for the annualised residual risk.

4.4. The Smart Alpha portfolio: does sparsity matter?

We may recall that the Smart Alpha portfolio is based on a machine learning method, the SPCA version of Wu and Chen (2016, see Section 3.2 above), that allows the sparse exposures of a stock and then its alpha to be estimated. In this section we evaluate the benefit of using SPCA rather than the traditional PCA. To do this, we consider the Smart Alpha portfolio optimisation programme in (24), with the systematic covariance and the alphas of the stocks estimated from a standard PCA with the optimal number of latent factors instead of the Sparse version of PCA, which is

$$\begin{cases} \widetilde{\omega} = \arg\min_{\omega} \omega' \, \widetilde{\Sigma}_{S,pca}\omega \\ u.c.\ \omega'\widetilde{\alpha}_{pca} \geq \epsilon, \omega \geq 0, \omega \leq \overline{\omega}, \omega' e = 1. \end{cases} \tag{27}$$

Comparing the performances of the two strategies in (24) and (27) will help in evaluating the relevance of estimating the systematic covariance matrix and the alphas, which are the two building blocks of the Smart Alpha strategy, using the sparse loadings of the stocks. The results of the backtesting exercise are displayed in table 6.

The results in table 6 show the absolute performance or raw return of the Smart Alpha portfolio based on the SPCA to be much higher than that of its equivalent based on the traditional PCA. Over the entire period, the performance of the SPCA Smart Alpha portfolio is equal to 584.01%, while that of the PCA version is 466.36% for instance. This dominance also holds when the focus is on the annualised average return, the annualised average excess return, the annualised alpha, the Appraisal ratio, or to a lesser extent the annualised volatility and the maximum drawdown.

It is worth mentioning that sparsity has an interesting effect on portfolio turnover. While the turnover of the SPCA Smart Alpha portfolio is equal to 20.56, the value of the PCA equivalent is higher at 22.15. This result is expected because sparsity removes noise from the estimation of the stock's loadings by setting small values that are non informative and erratic over time to exactly zero. This de-noising process has the advantage of producing stable exposures and alphas for the stocks, and so there is less portfolio rebalancing and lower transaction costs. Though this difference may seem small at first glance, it should be noted that the figures displayed relate on average to a monthly rebalancing, and that over a long period like that of our backtesting, the gain in terms of transaction costs may be high.

5. Economic value of Smart Alpha strategy and sensitivity analysis

5.1. Economic value

To calculate the value of the economic gains associated with our Smart Alpha strategy, this section uses the utility-based measure developed in Fleming *et al.* (2001, 2002). It is based on quadratic utility as an approximation of the true utility function of the investor and assumes that relative risk aversion (γ) is constant. Under these conditions, the average realised utility for a given portfolio ($\overline{U}_p$) is

$$\overline{U}_p(\cdot) = W_0 \left[\sum_{t=0}^{T-1} R_{p,t+1} - \frac{\gamma}{2(1+\gamma)} R^2_{p,t+1} \right], \tag{28}$$

where W_0 is the initial wealth and $R_{p,t+1} = \widetilde{\omega}_t' r_{t+1}$ are the returns on the portfolio (p). To measure the value of switching

Table 7. Willingness-to-pay for different strategies versus the Smart Alpha based on SPCA.

	SA-PCA	SC	LC	V	G	Q	MOM
$\gamma = 1$	1.13	0.42	6.56	6.38	5.80	3.86	2.14
$\gamma = 3$	1.10	0.97	6.68	6.63	5.88	3.78	2.24
$\gamma = 5$	1.10	1.17	6.72	6.72	5.90	3.75	2.28
$\gamma = 10$	1.09	1.35	6.76	6.80	5.93	3.73	2.31
	MV	E	RP	EW	MV-OPT	BAB	
$\gamma = 1$	5.02	6.04	3.13	3.35	2.26	2.47	
$\gamma = 3$	5.03	6.19	3.39	3.71	2.23	2.50	
$\gamma = 5$	5.03	6.24	3.48	3.83	2.23	2.52	
$\gamma = 10$	5.04	6.29	3.56	3.95	2.22	2.53	

Source: Bloomberg, daily data from 3 December 2001 to 28 May 2018. Computations by the authors. SA-PCA refers to our Smart Alpha portfolio strategy based on the traditional PCA, SC the MSCI Europe Small Capitalisation Total Return, LC the MSCI Europe Large Capitalisation Total Return, V the MSCI Europe Value Total Return, G the MSCI Europe Growth Total Return, Q the MSCI Europe Quality Total Return, MOM, the MSCI Europe Momentum Total Return, MV the MSCI Europe Minimum Volatility Total Return, E the MSCI Europe Index, RP the risk parity portfolio, EW the equally weighted portfolio, MV-OPT the optimised minimum volatility portfolio, and BAB the Betting-Against-Beta portfolio.

Table 8. OLS regressions of the Smart Alpha strategy on the five Fama-French European factors.

# Model	Intercept	Market	SMB	HML	RMW	CMA	*R*-squared
			Panel A: Univariate OLS regressions				
(1)	9.792% ***	0.447***					0.602
	(5.619)	(30.918)					
(2)	13.463%***		−0.279***				0.134
	(4.356)		(−9.611)				
(3)	12.852%***			0.073**			0.005
	(4.312)			(2.379)			
(4)	12.888%***				0.020		0.000
	(4.292)				(0.687)		
(5)	13.088%***					−0.0706*	0.005
	(4.358)					(−1.681)	
			Panel B: Multivariate OLS regressions				
(6)	9.757%***	0.450***	0.008				0.602
	(5.599)	(28.353)	(0.599)				
(7)	9.311%***	0.529***	0.140***	−0.228***			0.628
	(5.581)	(27.178)	(5.883)	(−6.159)			
(8)	9.228%***	0.506***	0.068	−0.240***	0.112***		0.634
	(5.564)	(21.281)	(1.610)	(−7.564)	(3.004)		
(9)	9.200%***	0.501***	0.074*	−0.195***	0.151***	−0.088*	0.635
	(5.563)	(20.887)	(1.943)	(−4.143)	(3.146)	(−1.811)	

Source: Bloomberg, Fama-French European five-factor model. Daily data from 3 December 2001 to 28 May 2018. Computations by the authors. The Newey-West corrected *t*-statistic of each parameter is displayed in parentheses below the estimated parameter. The $\bar{R}^2$ is the adjusted R^2. The intercept is the annualised value of the estimated portfolio's alpha. *, **, and *** denote traditional significance at the 10%, 5% and 1% levels respectively.

from a reference trading strategy to a candidate one, we then equate their average realised utilities,

$$\sum_{t=0}^{T-1}(R_{2,t+1} - \Delta) - \frac{\gamma}{2(1+\gamma)}(R_{2,t+1} - \Delta)^2 = \sum_{t=0}^{T-1} R_{1,t+1} - \frac{\gamma}{2(1+\gamma)} R_{1,t+1}^2, \tag{29}$$

where $R_{1,\,t+1}$ and $R_{2,t+1}$ are the returns for the two strategies. To equate the average utilities, we subtract a constant, Δ, from each of the returns on the candidate strategy. This represents the cost that the investor would be able to pay, say for performance fees, while still having the same expected utility as under the reference strategy.

Table 7 reports the annual fees in per cent that an investor with a quadratic utility function and constant relative risk aversion equal to γ would be willing to pay to switch from the SPCA Smart Alpha strategy to each of the alternative competitive strategies considered in the sections above. The fees in the table are determined empirically as the value that equalises the ex-post utility for the two different portfolios calculated from annual returns. The results show that all the values are positive, meaning that a rational investor switching from the SPCA Smart Alpha strategy to one of the alternative strategies has to pay some positive performance cost. To switch from the SPCA Smart Alpha strategy to its PCA analogue for instance, a rational agent with a quadratic utility function and a relative risk aversion coefficient $\gamma = 3$ must pay a cost equivalent to 1.10% per year over the full sample. It can be concluded from the table that all the alternative competing portfolios are inferior to the SPCA Smart Alpha portfolio.

5.2. *Sensitivity analysis*

In this last section, we analyse the sensitivity of the (SPCA) Smart Alpha portfolio to the five long-short based Fama-French (1993, 1996 and 2015) European factors.† The market factor is the return in euros on the European value-weighted market portfolio minus the US one month T-bill rate. The SMB (Small Minus Big) factor is the average return on the small stock portfolios minus the average return on the big stock portfolios. The HML (High Minus Low) factor is the average return on the value portfolios minus the average return on the growth portfolios. The RMW (Robust Minus Weak) factor is the average return on the robust operating profitability portfolios minus the average return on the weak operating profitability portfolios. The CMA (Conservative Minus Aggressive) factor is the average return on the conservative investment portfolios minus the average return on the aggressive investment portfolios.

To measure the sensitivity of the Smart Alpha strategy to factors, we performed univariate and multivariate OLS regressions on the five Fama-French European factors. Table 8 displays the results for the whole sample ranging from 3 December 2001 to 28 May 2018. In table 8, Panel A presents the estimated coefficients of the univariate OLS regressions and Panel B displays the estimated coefficients of the multivariate OLS regressions. We made a Newey-West correction (Newey and West 1987) for the inference to deal with heteroskedasticity and autocorrelation.

† The Fama-French five factors are constructed using the six value-weighted portfolios formed on size and book-to-market, the six value-weighted portfolios formed on size and operating profitability, and the six value-weighted portfolios formed on size and investment.

In Panel A of table 8, except for Model 1 which exhibits a moderately high R-squared, the other Fama-French factors (Model 2–Model 5) have weak explanatory powers with adjusted R-squared near 0 for HML, RMW and CMA factors. We observe that the explanatory powers increase when we move from Panel A to Panel B, but these increases are marginal. Indeed, the adjusted R-squared reaches only 0.635 for the five Fama-French Factors regression (Model 9), while the univariate model with the market factor has an adjusted R-squared of 0.602.† This is the proof that the market factor is the only factor that can explain the dynamics of the Smart Alpha portfolio strategy, but even then only weakly. The incremental value of the other four factors appears limited.

This last observation is reinforced when both the values of the t-statistics and the estimated alphas are considered. Indeed, results in Harvey *et al.* (2016) provide guidance about inference in factorial regression analysis. These authors stress the data mining problem arising from the many empirical papers that try to explain the cross-section of expected returns. They use a multiple testing approach to analyse the significance of a factor given the previous tests on other factors, and they argue that a t-statistic with an absolute value that is greater than 3, and so is above the usual 1% critical value of the Student distribution, should be used instead for inference. Following this route, we can observe that most of the estimated coefficients in table 8 are significantly different from 0 with a t-statistic higher than 3.0, except for HML, RMW and CMA in the univariate regressions. The estimated values for the alphas meanwhile are always positive and significant for all of the univariate and multivariate models in table 8. The annualised alphas are between 9.311% and 13.463%.

All these results confirm that the sparse PCA methodology used in conjunction with the estimation of the optimal number of factors from the information criterion provides accurate estimates of the alpha of a stock and its exposures to latent factors, and so leads to optimal portfolios that meet the idea underlying the Smart Alpha strategy: reducing the exposures to latent factors while generating alpha.

6. Conclusion

In this article, we propose an active investing approach to allocating equity portfolios. The strategy consists of maximising alpha, which is the expected return uncorrelated to various systematic sources of risk, while minimising the exposures to those same systematic risk factors. The core idea is to bet on alpha rather than on alternative risk premiums in the way that the factor investing approach does.

Methodologically, factors and the exposures and alphas of the stocks are extracted through a latent factor model estimated by Sparse Principal Component Analysis. In this framework, timing the exact number of relevant factors is crucial, and we achieve this by using the information criteria-like statistics in Bai and Ng (2002).

† For comparison, the same exercise for the Risk Parity portfolio considered above leads to an adjusted R-squared of 0.865 for model 9.

The empirical results confirm the value of our framework. The Smart Alpha portfolio has both lower risk and higher returns, alpha, than the market cap-weighted index, popular MSCI factor investing indexes, the Betting-Against-Beta strategy, and other heuristic or smart beta approaches such as the naïve equally weighted portfolio, the risk parity portfolio and the minimum variance portfolio. Economically, our strategy also appears to perform best, because a rational investor who switches from our strategy to one of the alternatives has to pay a positive and significant performance cost. These results disprove the current folklore in the asset management industry about the death of alpha and the superiority of factor investing.

Moreover, we evaluate the economic value of estimating the exposures and alphas of the stocks using Sparse Principal Component Analysis instead of the traditional Principal Component Analysis, and we observe significant economic gains. Lastly, attribution analyses show that our strategy delivers ex-post returns that are not explained by the traditional empirical factors. This result is in line with the philosophy underlying the strategy of minimising the portfolio's exposures to the main sources of risk.

We may finish by noting that, because our framework is general and not specific to the equity market, the Smart Alpha portfolio methodology can be adapted to other asset classes and also to portfolio managers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Ahn, S.C. and Horenstein, A.R., Eigenvalue ratio test for the number of factors. *Econometrica*, 2013, **81**(3), 1203–1227.
Ahn, S.C., Horenstein, A.R. and Wang, N., Beta matrix and common factors is stock returns. *J. Financ. Quant. Anal.*, 2018, **53**(3), 1417–1440.
Amihud, Y., Illiquidity and stock returns: Cross-section and time-series effects. *J. Financ. Mark.*, 2002, **5**(1), 31–56.
Bai, J. and Ng, S., Determining the number of factors in approximate factor models. *Econometrica*, 2002, **70**(1), 191–221.
Bai, J. and Ng, S., Evaluating latent and observed factors in macroeconomics and finance. *J. Econom.*, 2006, **131**, 507–537.
Bai, J. and Ng, S., Determining the number of primitive shocks in factor models. *J. Bus. Econ. Stat.*, 2007, **25**(1), 52–60.
Behr, P., Guettler, A. and Truebenbach, F., Using industry momentum to improve portfolio performance. *J. Bank. Financ.*, 2012, **36**(5), 1414–1423.
Bloomfield, T., Leftwich, R. and Long, J.B., Portfolio strategies and performance. *J. Financ. Con.*, 1977, **5**(2), 201–218.
Cadima, J. and Jolliffe, I.T., Loadings and correlations in the interpretation of principal components. *J. Appl. Stat.*, 1995, **22**, 203–214.
Calomiris, C.W., Love, I. and Peria, M.S.M., Stock returns' sensitivities to crisis shocks: Evidence from developed and emerging markets. *J. Int. Money Finance*, 2012, **31**(4), 743–765.
Choueifaty, Y. and Coignard, Y., Towards maximum diversification. *J. Portfolio Manage.*, 2008, **35**(1), 40–51.
Cochrane, J.H., Presidential address: Discount rates. *J. Finance*, 2011, **66**(4), 1047–1108.

Connor, G. and Korajczyk, R.A., A test for the number of factors in an approximate factor model. *J. Finance*, 1993, **48**(4), 1263–1291.
d'Aspremont, A., El Ghaoui, L., Jordan, M.I. and Lanckriet, G.R.G., A direct formulation for sparse PCA using semidefinite programming. *SIAM Rev.*, 2007, **49**(3), 434–448.
de Bondt, W.F.M. and Thaler, R.H., Does the stock market overreact?. *J. Finance*, 1985, **40**(3), 793–805.
de Bondt, W.F.M. and Thaler, R.H., Further evidence on investor overreaction and stock market seasonality. *J. Finance*, 1987, **42**(3), 557–581.
DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Rev. Financ. Stud.*, 2009, **22**, 1915–1953.
Donald, S.G., Inference concerning the number of factors in a multivariate nonparametric relationship. *Econometrica*, 1997, **65**(1), 103–132.
Fama, E.F. and French, K.R., The cross-section of expected stock returns. *J. Finance*, 1992, **47**(2), 427–465.
Farcomeni, A., An exact approach to sparse principal component analysis. *Comput. Stat.*, 2009, **24**(4), 583–604.
Feng, C.-M., Gao, Y.-L., Lui, J.-X., Zheng, C.-H., Li, S.-J. and Wang, D. (Eds.), *A Simple Review of Sparse Principal Components Analysis*, Lecture Notes in Computer Science, 2016 (Springer: Cham).
Fleming, J., Kirby, C. and Ostdiek, B., The economic value of volatility timing. *J. Finance*, 2001, **56**(1), 329–352.
Fleming, J., Kirby, C. and Ostdiek, B., The economic value of volatility timing using 'realized' volatility. *J. Financ. Econ.*, 2002, **67**(3), 473–509.
Frazzini, A. and Pedersen, L.H., Betting against beta. *J. Financ. Econ.*, 2014, **111**(1), 1–25.
Gennaioli, N. and Shleifer, A., *A Crisis of Beliefs: Investor Psychology and Financial Fragility*, 2018 (Princeton University Press: Princeton, NJ).
Geweke, J., The dynamic factor analysis of economic time series. In *Latent Variables in Socioeconomic Models*, edited by I.D. Aigner and A. Goldberger, pp. 365–383, 1977 (North-Holland: Amsterdam).
Harvey, C.R., Liu, Y. and Zhu, H., $\cdots$. and the cross-section of expected returns. *Rev. Financ. Stud.*, 2016, **29**(1), 5–68.
Jagannathan, R. and Ma, T., Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance*, 2003, **58**(4), 1651–1683.
Jegadeesh, N. and Titman, S., Returns to buying winners and selling losers: Implications for stock market efficiency. *J. Finance*, 1993, **48**(1), 65–91.
Jolliffe, I.T., Trendafilov, N.T. and Uddin, M., A modified principal component technique based on the lasso. *J. Comput. Graph. Stat.*, 2003, **12**(3), 531–547.
Journée, M., Nesterov, Y., Richtarik, P. and Sepulchre, R., Generalized power method for sparse principal component analysis. *J. Mach. Learn. Res.*, 2010, **11**, 517–553.
Kelly, T.B., Pruitt, S. and Su, Y., Characteristics are covariances: A unified model of risk and return. *J. Financ. Econ.*, 2019, **134**(3), 501–524.
Kourtis, A., Dotsis, G. and Markellos, R.N., Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. *J. Banking Finance*, 2012, **36**(9), 2522–2531.
Ledoit, O. and Wolf, M., Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Financ.*, 2003, **10**(5), 603–621.
Lee, C.M.C., Shleifer, A. and Thaler, R.H., Investor sentiment and the closed end fund puzzle. *J. Finance*, 1991, **46**(1), 75–109.
Lewbel, A., The rank of demand systems: Theory and nonparametric estimation. *Econometrica*, 1991, **59**(3), 711–730.
Lewellen, J., Nagel, S. and Shanken, J., A skeptical appraisal of asset pricing tests. *J. Financ. Econ.*, 2010, **96**(2), 175–194.
Maillard, S., Roncalli, T. and Teiletche, J., The properties of equally weighted risk contribution portfolios. *J. Portfolio Manage.*, 2010, **36**(4), 60–70.
Markowitz, H., Portfolio selection. *J. Finance*, 1952, **7**(1), 77–91.
Meucci, A., Managing diversification. *Risk*, 2009, 74–79.
Newey, W.K. and West, K.D., A simple, positive semidefinite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 1987, **55**(3), 703–708.
Onatski, A., Determining the number of factors from empirical distribution of eigenvalues. *Rev. Econ. Stat.*, 2010, **92**(4), 1004–1016.
Sargent, T.J. and Sims, C.A., Business cycle modeling without pretending to have too much a priori economic theory. *New Methods Business Cycle Res.*, 1977, **1**, 145–168.
Sharpe, W.F., Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance*, 1964, **19**(3), 425–442.
Shiller, R.J., Narrative economics. *Am. Econ. Rev.*, 2017, **107**(4), 967–1004.
Stock, J.H. and Watson, M.W., Forecasting inflation. *J. Monet. Econ.*, 1999, **44**(2), 293–335.
Tibshirani, R., Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B*, 1996, **58**, 267–288.
Treynor, J. L., How to rate management of investment funds. *Har. Bus. Rev.*, 1965, **43**(1), 63–75.
Tu, J. and Zhou, G., Markowitz meets talmud: A combination of sophisticated and naïve diversification strategies. *J. Financ. Econ.*, 2011, **99**(1), 204–215.
Wang, Y. and Wu, Q., Sparse PCA by iterative elimination algorithm. *Adv. Comput. Math.*, 2012, **36**(1), 137–151.
Wu, M.-C. and Chen, K.-C., Sparse PCA via hard thresholding for blind source separation. IEEE International Conference on Acoustics, Speech and Signal Processing, 2016.
Zaremba, A., Umutlu, M. and Karathanasopoulos, A., Alpha momentum and alpha reversal in country and industry equity indexes. *J. Empir. Financ.*, 2019, **53**(C), 144–161.
Zou, H. and Hastie, T., Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B*, 2005, **67**, 301–320.
Zou, H., Hastie, T. and Tibshirani, R., Sparse principal component analysis. *J. Comput. Graph. Stat.*, 2006, **15**(2), 265–286.

Appendix A. Sparse PCA algorithm

This Appendix presents the iterative SPCA algorithm of Wu and Chen (2016). In our context, the algorithm solves for the constrained penalised regression in (17) or equivalently (18) by alternately fixing one of the arguments A or B, and solving in the dimension of the free argument B or A. The normalised components $\widetilde{b}_j, j = 1, \ldots, \widehat{m}$ of $\widetilde{B}$ are the sparse loadings of the stocks, with $\widehat{m}$ as the number of factors.

- Step 1. Let $A = [a_1, \ldots, a_{\widehat{m}}]$ start at $V[:, 1 : \widehat{m}]$ with V as the stock's PCA loadings on the first $\widehat{m}$ principal components, obtained from the singular value decomposition (SVD) of R (see equation (14)).
- Step 2. Given a fixed A, apply the hard-thresholding rule in (20) to get sparse loadings, for $j = 1, \ldots, \widehat{m}$, with

$$\widetilde{b}_j = D_j R' R \widetilde{a}_j. \tag{A1}$$

- Step 3. For a fixed $\widetilde{B} = [\widetilde{b}_1, \ldots, \widetilde{b}_{\widehat{m}}]$, compute the SVD of $X'X\widetilde{B} = \widetilde{U}\widetilde{\Omega}\widetilde{V}'$, then update $A = \widetilde{U}\widetilde{V}'$.
- Step 4. Repeat Steps 2–3, until convergence.
- Setp 5. Get sparse PCA loadings via normalisation, with $\widetilde{\Lambda}_{spca} = [\widetilde{\lambda}_{1,spca}, \ldots, \widetilde{\lambda}_{\widehat{m},spca}]$, and

$$\widetilde{\lambda}_{j,spca} = \frac{\widetilde{b}_j}{\|\widetilde{b}_j\|_2}, \quad j = 1, \ldots, \widehat{m}. \tag{A2}$$

Appendix B. Additional table

Table B1. Latent and a-priori factor matching (absolute value of the correlation coefficients).

Number of latent factors	1	2	3	4	5
Calendar years					
2001	Rm	Sect_Media	Sect_Travel		
	1.00	0.65	0.52		
2002	Rm	Sect_Tech	Sect_Media	Sect_Oil	Sect_Health
	1.00	0.44	0.40	0.32	0.29
2003	Rm	LARGE	Sect_Chem		
	1.00	0.58	0.28		
2004	Rm	Sect_Media			
	1.00	0.49			
2005	Rm	SMALL	Sect_Basic		
	1.00	0.20	0.42		
2006	Rm	SMALL			
	1.00	0.75			
2007	Rm	Sect_Travel			
	1.00	0.51			
2008	Rm	LARGE	Sect_Bks	VALUE	GROWTH
	1.00	0.71	0.72	0.24	0.39
2009	Rm	Sect_Bks	SMALL	Sect_Ins	QUALITY
	1.00	0.48	0.64	0.60	0.39
2010	Rm	Sect_Bks	VALUE	SMALL	Sect_Chem
	1.00	0.52	0.69	0.29	0.36
2011	Rm	Sect_Bks	LARGE		
	1.00	0.76	0.58		
2012	Rm	Sect_Bks	VALUE	LARGE	Sect_Ins
	1.00	0.68	0.73	0.49	0.16
2013	Rm	Sect_Travel	VALUE		
	1.00	0.72	0.72		
2014	Rm	Sect_Bks	Sect_Oil	MOMENTUM	
	1.00	0.75	0.72	0.32	
2015	Rm	Sect_Oil	VALUE	QUALITY	Sect_Bks
	1.00	0.83	0.16	0.61	0.43
2016	Rm	Sect_Basic	Sect_Bks	Sect_Travel	QUALITY
	1.00	0.87	0.78	0.75	0.56
2017	Rm	Sect_Travel	Sect_Tech	Sect_Oil	MINIMUM VARIANCE
	1.00	0.16	0.66	0.47	0.58
2018	Rm	Sect_Tech	SMALL	Sect_Bks	
	1.00	0.51	0.21	0.77	

Source: Bloomberg, DataStream; daily data from 30 November 2001 to 28 May 2018. Computations by the authors. Rm: MSCI Europe; VALUE: MSCI Europe Value; LARGE: MSCI Europe Large; SMALL: MSCI Europe Small; MOMENTUM: MSCI Europe Momentum; QUALITY: MSCI Europe Quality; MINIMUM VARIANCE: MSCI Minimum Variance; Sect_Basic: DS Equity Basic Resources; Sect_Bks: DS Equity Banks; Sect_Chem: DS Equity Chemistry; Sect_Oil: DS Equity Oil & Gas; Sect_Tech: DS Equity Technology; Sect_Travel: DS Equity Travel & Leisures; Sect_Media: DS Media; Sect_Ins: DS Insurance; Sect_Health: DS Health.

Table B2. Sparsity degree of estimated latent factors.

Number of latent factors	1	2	3	4	5
Calendar years					
2001	Rm	Sect_Media	Sect_Travel		
	0	0.64	0.78		
2002	Rm	Sect_Tech	Sect_Media	Sect_Oil	Sect_Health
	0	0.77	0.82	0.89	0.89
2003	Rm	LARGE	Sect_Chem		
	0	0.71	0.84		
2004	Rm	Sect_Media			
	0	0.87			
2005	Rm	SMALL	Sect_Basic		
	0	0.92	0.88		
2006	Rm	SMALL			
	0	0.82			
2007	Rm	Sect_Travel			
	0	0.71			
2008	Rm	LARGE	Sect_Bks	VALUE	GROWTH
	0	0.54	0.59	0.63	0.64
2009	Rm	Sect_Bks	SMALL	Sect_Ins	QUALITY
	0	0.67	0.69	0.71	0.75
2010	Rm	Sect_Bks	VALUE	SMALL	Sect_Chem
	0	0.66	0.67	0.83	0.90
2011	Rm	Sect_Bks	LARGE		
	0	0.74	0.91		
2012	Rm	Sect_Bks	VALUE	LARGE	Sect_Ins
	0	0.71	0.82	0.84	0.91
2013	Rm	Sect_Travel	VALUE		
	0	0.84	0.87		
2014	Rm	Sect_Bks	Sect_Oil	MOMENTUM	
	0	0.79	0.81	0.81	
2015	Rm	Sect_Oil	VALUE	QUALITY	Sect_Bks
	0	0.49	0.87	0.51	0.92
2016	Rm	Sect_Basic	Sect_Bks	Sect_Travel	QUALITY
	0	0.51	0.53	0.46	0.75
2017	Rm	Sect_Travel	Sect_Tech	Sect_Oil	MINIMUM VARIANCE
	0	0.86	0.82	0.87	0.91
2018	Rm	Sect_Tech	SMALL	Sect_Bks	
	0	0.80	0.90	0.91	

Source: Bloomberg, DataStream; daily data from 30 November 2001 to 28 May 2018. Computations by the authors. Rm: MSCI Europe; VALUE: MSCI Europe Value; LARGE: MSCI Europe Large; SMALL: MSCI Europe Small; MOMENTUM: MSCI Europe Momentum; QUALITY: MSCI Europe Quality; MINIMUM VARIANCE: MSCI Minimum Variance; Sect_Basic: DS Equity Basic Resources; Sect_Bks: DS Equity Banks; Sect_Chem: DS Equity Chemistry; Sect_Oil: DS Equity Oil & Gas; Sect_Tech: DS Equity Technology; Sect_Travel: DS Equity Travel & Leisures; Sect_Media: DS Media; Sect_Ins: DS Insurance; Sect_Health: DS Health.