A factor-based risk model for multi-factor investment strategies

F. Abergel*, B. Bellone, F. Soupé

Quantitative Research Group
BNP Paribas Asset Management

July 7, 2021

Abstract

This note presents a novel, practical approach to risk management for multi-factor equity investment strategies. Our approach lies in the construction of a cross-sectional risk model using the stock return betas and a small number of style factors and macro-sectors indicator functions as explanatory variables in a cross-sectional regression. The model leads to a covariance structure that incorporates in an intuitive fashion the stocks' characteristics while at the same time possesses good conditioning properties leading to a robust optimization problems. Various portfolio constructions are analyzed in details, and some concrete examples are provided.

1 Introduction

Risk modelling is an area that lies at the very core of the financial industry. In a broad sense, it pertains to the understanding, statistical analysis and mathematical modelling of the dependencies between tradeable assets. For instance, the modelling of variance-covariance matrices, of tail risks, of credit risk, the design of stochastic volatility and correlation models... are intrinsically part of risk modelling. This article addresses the risk management of multi-factor investment strategies. Much as any investment strategy, a multi-factor investment portfolio is built using a risk model that is fully determined, in the usual quadratic framework, by a covariance structure between stock returns. There is however a host of different ways to define such a covariance structure: one can rely on empirical models, time series based-factor models, cross-sectional models... Our aim in this paper is to design a risk model that is specific to the context of **factor investing**, by providing a simple, inutitive and robust covariance structure that faithfully reflects the stock characteristics.

Let us first briefly recall what we mean by "factors".

*frederic.abergel@bnpparibas.com

Factors are characteristics of assets that are important to explain their risk and performance. For instance, in the stock market, the Capital Asset Pricing Model (CAPM) model asserts that the performance of a stock should be determined by a single stock characteristic, the **beta**. However, it is now widely accepted, and vastly documented in the academic literature, see e.g. the seminal papers [5] and [10], that other factors known as Value, Low Volatility, Quality and Momentum play a role in explaining stock returns.

Value characteristics such as the price-to-book or the price-to-earnings of a company measure the relative cheapness of a stock. It has been found that on average, over time, cheaper stocks tend to out-perform other stocks, in particular expensive stocks. **Volatility** is a measure of a stock price fluctuations over time. Evidence that less volatile stocks generate at least comparable returns to riskier stocks renders stocks with lower volatility more attractive for investors: same returns in the medium to long-term with less uncertainty. It is also known that higher **quality** stocks, e.g. the most profitable companies, tend to generate higher returns than other stocks, in particular when compared to the least profitable companies. Different measures of profitability can be used, e.g. return-on-equity. Last, stocks with the strongest price trends, e.g. stocks with the strongest out-performance relative to other stocks as measured over the previous 12 months also tend to continue to outperform. This is known as the **momentum** effect.

The building blocks of factor investing are the **single factor portfolios** built from suitable ranking, aggregation, normalization and neutralization. These are long-short portfolios with weights that reflects the score of an asset according to one particular member of the factor family. From there on, **multi-factor investing** consists in building, based on these single factor portfolios, a multi-factor, investable portfolio. This rather involved industrial process can be decomposed into two major steps. The first one requires generating a global score for each stock, leading to a theoretical long-short portfolio. The second, more technical step amounts to building an investment portfolio meeting a set of requirements such as weight positivity, diversification, low turnover or controlled volatility.

Let us now turn towards the main topic of this paper, namely, the design of a **risk model**.

A risk model consists in the design of the dependency structure between assets and the analysis of its evolution over time. It is an essential tool for the construction of an investment portfolio, which in its simplest form due to Markowicz relies on two components: a set of expected returns and a covariance matrix. In principle, one could design separately these two components of the optimization program. It is however useful and important, as analyzed and explained in for instance [13], that the risk model, as a portfolio construction tool, be consistent with the view one has on the market, that is, with the set of expected returns. The focus throughout this paper is on portfolio construction for multi-factorial investment strategies. Since such strategies typically use the factor scores as the building blocks of their expected returns, we will therefore concentrate on factor-based risk models. Specifically, we shall adopt a cross-sectional view to explain the stock return "alphas" - the excess returns over the

CAPM returns - by the characteristic portfolio weights and some additional, industrial sector-based variables.

The paper is organized as follows: Section 2 introduces some notations, Section 3 describes the general modelling approach and Section 4 studies the main object of interest, the covariance matrix. Section 5 analyzes the building blocks of a long-only investment portfolio based on this risk model, while Section 6 presents and discusses some concrete applications in the multi-factor equity universe.

2 Some notations

We introduce some mathematical notations to be used in the rest of this article. **Bold face** letters stand for vectors or matrices.

- N the total number of assets in the universe
- $\mathbf{R}$ the vector of stock returns in excess of cash
- $\boldsymbol{\beta}$ the vector of stocks' betas w.r. to the market returns
- $\mathbf{I}$ the vector of benchmark weights
- RM the benchmark (market) return $RM = \mathbf{I}.\mathbf{R}$
- $\mathbf{W}$ the investment portfolio weights
- β_W the investment portfolio beta $\beta_w \equiv \mathbf{W}.\boldsymbol{\beta}$
- $\mathbf{Q}$ the investment portfolio active weights $\mathbf{Q} = \mathbf{W} - \mathbf{I}$
- $\mathbf{z}_k$, $k = 1, ..., K$ the factor long-short portfolio weights
- $\mathbf{z}_k$, $k = K+1, ..., K+S$ the sector indicatrix portfolio weights
- for any vector $\mathbf{X}$, $\mathbf{X}^+$ stands for the positive part of $\mathbf{X}$ with components $\mathbf{X}_i^+ \equiv \mathbf{X}_i \mathbf{1}_{\{\mathbf{X}_i > 0\}}$

3 Designing a factor-based risk model

This section describes the general features of the factor-based risk model under scrutiny in this paper. By nature, the risk model and portfolio construction analyzed in this paper are essentially model for the **alpha** of stocks and portfolios, a fact that is already reflected in the very construction of the characteristic portfolios as long-short, low beta portfolios. Eventually, we will design a long-only, investment portfolio with a beta equal to 1, and hope that its excess returns over the benchmark yield some good performances. As a consequence, it is natural to deal separately with the beta and alpha part of the model, as we now explain.

Consider then the simple CAPM representation

$$R_i(t) = \beta_i RM(t) + \alpha_i. \tag{1}$$

The plan is the following: first, the market beta of the stocks is computed based on **time-series** regression; then, a **cross-sectional** regression is used to explain the stock return alphas. This dual approach is rather classical and has been widely advertised and studied, see e.g. [11][8][4]

for general mixed models of this type. However, let us point out right away that the specific model we have in mind is both parsimonious and adapted to multi-factor investing, so that the risk factors will pretty much be the same factors as those used to model the stock expected returns. This point will be developped at greater length in Section 5.

3.1 The stock betas

The first task at hand is of course to identify the stock betas. Broadly speaking, for each asset in the investment universe, the time series of its returns are linearly regressed against that of the market returns. There are several specific choices to be made, pertaining to the time horizon, weighing method, reference index... and we do not go into such details here. Suffice it to say that, from now on, the vector of stock betas is assumed to be available at any time *via* a systematic regression method.

3.2 Factor- and sector-based cross-sectional model

Once the betas are determined, the assets' alphas[1] are explained *via* a cross-sectional regression of the returns against some explanatory variables that have a direct interpretation as elementary portfolio weights. Such portfolios are of two types: long-short portfolios corresponding to the style factors Low volatility, Momentum, Quality and Value, and pure indicator portfolios corresponding to macro-sectors. Note that there are many additional evolutions that could be made: one can add "fuzziness" to the macro-sector definition, allowing portfolio weights to sit between 0 and 1 instead of being equal to 0 or 1. One can also incorporate other factors, for instance a Size factor, or enrich the sector portfolios with country or region indicators. The methodology we present here is quite general, and we have simply restricted the analysis to some specific factors so as to provide concrete examples.

In a continuous-time framework, a general factor model for the stock return alphas with sectors and style factors can be written as follows:

$$\begin{aligned} dX_i(t) &\equiv \frac{dS_i(t)}{S_i(t)} - \beta_i(t)dM(t) \\ &= \sum_{k=1}^{K} z_{ki}(t)d\lambda_k(t) + \sum_{k=K+1}^{K+S} z_{ki}(t)d\lambda_k(t) + d\epsilon_i(t) \end{aligned}$$

or, using vector notations:

$$d\boldsymbol{X}(t) = \sum_{k=1}^{K} \mathbf{z}_k(t)d\lambda_k(t) + \sum_{k=K+1}^{K+S} \mathbf{z}_k(t)d\lambda_k(t) + d\boldsymbol{\epsilon}(t). \qquad (2)$$

The stochastic processes $\lambda_k(t)$ involved in Equation (2) represent the (model-based) excess returns over the benchmark of the characteristic $(k = 1, ..., K)$ and sector indicator $(k = K+1, ..., K+S)$ portfolios, see the notations in Section 2.

[1] 'alpha 'will always refer to the CAPM alpha introduced in Equation (1)

In practice, we consider a discrete-time version of (2) for monthly returns

$$\boldsymbol{X} = \sum_{k=1}^{K+S} \mathbf{z}_k \lambda_k + \boldsymbol{\epsilon} \tag{3}$$

where the K first vectors $\mathbf{z}_k$ and factor returns λ_k correspond to the style factors, and the S last, to macro-sectors.

The model-implied approximate returns λ_k simply obtain as the linear regression solution to the approximate equation

$$\widetilde{\boldsymbol{X}} = \sum_{k=1}^{K+S} \mathbf{z}_k \lambda_k. \tag{4}$$

Once Equation (4) is solved, one can then undertake the task of analyzing the covariance structure it implies for the vector of stock return alphas $\mathbf{X}$. Note that in the rest of the paper, the term "covariance matrix", "covariance structure", "risk model"... will always implicitly refer to the corresponding object obtained from stock return alphas.

Remark In the examples we analyze, the style factor portfolios have been market- and sector-neutralized:

$$\begin{aligned} \forall (p,k) \in \{1,...,K\} \times \{K+1,...,K+S\}, \mathbf{z}_p.\mathbf{z}_k &= 0 \\ \forall p \in \{1,...,K\}, \mathbf{z}_p.\boldsymbol{\beta} &= 0, \end{aligned}$$

and similar equalities obviously hold between the sector portfolios themselves. As a consequence, the OLS solution to Equation (4) can be decomposed into two independent factor and sector regressions. Between factors, we do not enforce any additional orthogonality conditions, as we prefer to retain as much as possible the financial interpretation of each style factor. For a more detailed analysis of the cross-sectional dependency structure between the factor portfolios, we refer the interested reader to [1].

3.3 The cross-sectional factor returns and portfolios

The cross-sectional approach paves the way to a new understanding of the single factor portfolios, as we now explain. Consider again Equation (4) and introduce the Gram matrix $\mathbf{G}$ of the factor portfolios with entries

$$\mathbf{G}_{pq} \equiv \mathbf{z}_p.\mathbf{z}_q \; 1 \leqslant p, q \leqslant K.$$

Then, see the remark in Section 3.2 above, the solution to the factor part of the OLS regression in (3) is simply

$$\boldsymbol{\Lambda} = \mathbf{G}^{-1}\mathbf{RLS} \tag{5}$$

where $\mathbf{RLS}$ is the vector of realized returns of the original single factor portfolios : $\mathbf{RLS}_k \equiv \mathbf{z}_k.\mathbf{R}$.

The solution λ_k to Equation (5) are quite naturally termed the **cross-sectional** returns of the factor portfolios, and they also offer another intuitive interpretation: they can be viewed as the **realized** returns of modified factor portfolios $\mathbf{y}_k = {\mathbf{z}_k}^t(\mathbf{G}^{-1}) \equiv \mathbf{z}_k\mathbf{G}^{-1}$ ($\mathbf{G}$ is symmetric). We call these new portfolios $\mathbf{y}_k$ the **cross-sectional factor portfolios**. As we shall see in see Section 6.1 where their performances are analyzed, these cross-sectional factor portfolios have returns that are only weakly correlated over time - in any case, much less so than those of the original factor portfolios. This weak correlation property plays an important role in the good behaviour of the risk model and its covariance matrix.

4 Variance-covariance structure of the cross-sectional risk model

In a Gaussian framework where only quadratic risk is taken into account, the key feature of a risk model is the structure of its covariance matrix, and we now proceed to build it.

4.1 Building the alpha covariance matrix

Starting from the model specified in Equation (3), we posit that the covariance matrix should have two components: first, a term stemming form the "explained" part of the cross-sectional regression, based on the variance-covariance matrix of the cross-sectional returns in Equation (4). Then, a fully diagonal idiosyncratic term given by the variance of the "unexplained" part of the time series $\epsilon_i(t)$.

The variance-covariance structure associated to Equation (3) is therefore given by its two components:

- the factor and sector-based covariance matrix with entries $\mathbf{F}_{ij} \equiv \sum_{p,k=1}^{K+S} z_{ki}z_{pj}Cov(\lambda_k, \lambda_p)$;
- the idiosyncratic, diagonal matrix with entries $\mathbf{D}_{ii} = Var(\boldsymbol{\epsilon}_i)$;

and the global variance-covariance matrix $\mathbf{M}$ is simply defined by

$$\mathbf{M}_{ij} = \mathbf{F}_{ij} + \mathbf{D}_{ij}. \tag{6}$$

Note that the factor part $\mathbf{F}_{ij}$ in Equation (6) is close to the empirical covariance matrix of the cross-sectional factor portfolio returns, but the equality is not perfect since the weight vectors $\boldsymbol{z_k}$ change over time.

It is quite likely that this model fails to faithfully represent the real asset dynamics: the underlying assumption that the time series of error terms have no autocorrelation and be uncorrelated with the factor and sector returns is far from being guaranteed by the cross-sectional orthogonality property. We will however show that, simple as it is from an econometric point of view, see Section 4.2 below, this model is quite valuable when used as a risk model for portfolio construction.

A remark In practice, the empirical behaviour of the cross-sectional return covariance matrix with entries $Cov(\lambda_k, \lambda_p)$ suggests to make the convenient choice of a pure **diagonal** factor-only covariance matrix by setting the factor-factor correlations to 0.

4.2 The risk model from an econometric perspective

This short section is devoted to the assessment of the risk model from an econometric perspective. Quite generally speaking, this analysis is related to the rather deep question of understanding whether factors should be interpreted as *trend* or *risk* factors. From a mathematical and statistical point of view, this question amounts to determining whether the factor and sector returns λ, μ, when viewed as stochastic processes, have zero or finite quadratic covariations. There exists a host of theoretical results in the statistical theory of stochastic processes providing tools to answer this question [2], but a practical answer is hard to get in the situations we consider, because the cross-sectional factor portfolio returns can only be sampled at very low frequency. Note that, to the contrary, the sector cross-sectional returns can be computed at a frequency as high as daily, making the analysis much more reliable.

Without dwelling too much on this theoretical difficulty, we can however present some results that shed a light on the econometric properties of Model (2), and the very important role played by the sector indicatrices. A rather straightforward numerical experiment has been performed, the idea being to generate Monte Carlo paths according to the dynamics of Model (2). It proceeds as follows:

- cross-sectional distributional assumptions are made on the factor portfolio weights, and they are calibrated onto the data;
- the cross-sectional factor and sector portfolio returns are simulated as a low-dimensional stochastic process with drift vector and covariation matrix calibrated onto the data;
- idiosyncratic risks are simulated as an N-dimensional, uncorrelated Wiener process.

Once the model inputs are calibrated onto the data, the model-implied correlation structure is computed from simulated paths[3]. The first natural question one may ask is then: does the covariance structure of the factor- and sector-based model resemble the empirical one ? Figure 1 shows the empirical joint distribution of realized and model-implied correlation for stocks in the $S\&P500$ universe, bringing a partially negative answer to that question. One can clearly see that the explanatory power of the model is rather limited and that it is necessary to incorporate the sector indicatrices in order to generate a correlation distribution possessing at least qualitatively, some of the features of the empirical distribution.

[2]the underlying idea behind those tools is the a.s. convergence of quadratic sums towards the integrated quadratic variation for a process with continuous sample paths, and the scaling properties of such sums as the frequency goes to 0

[3]It could actually be computed analytically, but we found it simpler and more flexible to

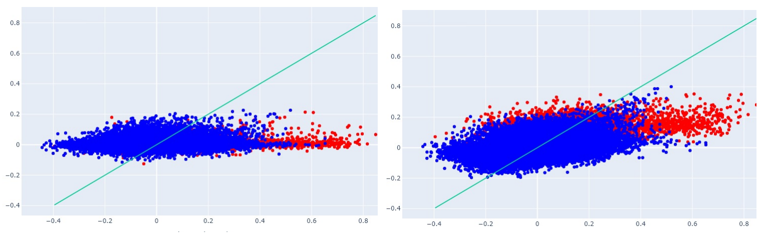

Figure 1: Model-based *vs* realized correlation without (left) or with (right) sectors

What is clear from this simple numerical exercise is that the assumption of zero correlation between the idiosyncratic risks is not realistic. This rather poor approximation of the empirical correlation between stocks prevents us from using Model (2) as a risk predictor, or even a risk management tool for general portfolios. This was expected, as the model only has a very low number of degree of freedom, and we shall see that its main usefulness lies rather in the great portfolio constrution tool it provides.

5 Portfolio construction

In this section, we revisit the portfolio construction problem. Although we shall eventually be using the cross-sectional risk model (4), some of the results we present here are rather general, and not at all specific to the cross-sectional risk model. We nonetheless think it worthwhile to revisit the portfolio construction process in details.

5.1 Building a multi-factor, theoretical long-short portfolio

Based on the four style factors Low volatility, Momentum, Quality and Value, there are of course infinitely many possibilities to build a theoretical, multi-factorial long-short portfolio. A natural and, by now, well-documented approach is the risk budgeting approach as analysed in [3], see also e.g. [7] for a more recent analysis of risk budgeting approaches. This risk-budgeting can be based on the original, single factor portfolios considered as elmementary tradeable assets. However, in the specific context under scrutiny here, the low correlation of the cross-sectional returns[4] suggest that we rather use the new cross-sectional factor portfolios introduced in Section 3.3.

For the sake of the argument, consider the equal risk budget (ERB) or contribution (ERC) portfolios, see e.g. [12], based on either the original or the cross-sectional long-short portfolio: one thus obtains four natural candidate multi-factor portfolios. When using the *cross-sectional* single

use Monte Carlo paths

[4]see Section 6.1 for details

factor portfolios, both approaches yield almost identical results, thanks to the low factor-factor correlation. As opposed to this case, ERB or ERC constructions based on the original single factor portfolio lead to quite different portfolios. Interestingly enough, the results in Section 6.3 show very little differences, in terms of performances, between the four candidate portfolios. Therefore, it makes good sense to use the simpler ERB approach based on the cross-sectional portfolios, and this will be the focus for the rest of this paper.

We will now denote by $\mathbf{A}$ the multi-factor portfolio weights: $\mathbf{A} = \sum_{i=1}^{K} a_k \mathbf{z}_k$ for a set of time-varying weights a_k, $k = 1, ..., K$ obtained *via* equal risk budgeting of the cross-sectional factor portfolios.

5.2 Building a long-only, investable portfolio

Setting aside for the sake of clarity the turnover and transaction cost constraints, we focus on the rather general case where the optimal portfolio $\mathbf{W}$ and its active weight vector $\mathbf{Q}$ can be written as the solution to a Markowitz portfolio optimization with the alpha variance-covariance matrix $\mathbf{M}$ and a set of expected returns $\boldsymbol{\mu}$:

$$
\begin{array}{rcl}
\sup & & \mathbf{q}.\boldsymbol{\mu} \\
\boldsymbol{q} & = & \boldsymbol{w} - \boldsymbol{I} \\
0 & \leqslant & w_i \\
q_i & \leqslant & q_{max} \\
\mathbf{q}.\boldsymbol{e} & = & 0 \\
\mathbf{M}\boldsymbol{q}.\boldsymbol{q} & \leqslant & TE^2.
\end{array}
\tag{7}
$$

In Problem (7), the objective function is naturally expressed in terms of the **active weight q**, as are most of the constraints. In fact, only the nonnegativity constraint involves the original weight vector $\mathbf{w}$.

There are obviously many different ways to introduce a set of expected returns derived from the factor portfolios. Here, the focus is on a special case - and one that is very important in practice - where the expected returns $\boldsymbol{\mu}$ come from an explicit transformation of the theoretical long-short portfolio.

Reverse optimization and portfolio construction The portfolio construction generating the investment strategies presented in this article rests on the useful and intuitive concept of reverse optimization as introduced by [9], see e.g. [2] for an in-depth analysis. When specialized to the case of multi-factor investing, the key underlying assumption is that the theoretical, multi-factorial long-short portfolio is, up to a multiplicative constant, the active weight vector solution to the Markowitz optimization probleme with the sole risk constraint:

$$
\begin{array}{rcl}
\sup & & \mathbf{q}.\boldsymbol{\mu} \\
\mathbf{M}\boldsymbol{q}.\boldsymbol{q} & \leqslant & TE^2,
\end{array}
\tag{8}
$$

the solution of which is

$$\mathbf{A} = C\mathbf{M}^{-1}\boldsymbol{\mu}$$

for some TE-related constant C.

This process is termed the *reverse optimization* because it builds a set of expected returns from a candidate portfolio rather than the usual other way around.

Since Problem (7) is invariant by multiplicative scaling of the expected returns, one simply imposes the relation

$$\boldsymbol{\mu} = \mathbf{MA},$$

and Problem (7) becomes

$$\begin{aligned} \sup \quad & \mathbf{q}.\mathbf{MA} \\ \boldsymbol{q} &= \boldsymbol{w} - \boldsymbol{I} \\ 0 &\leqslant w_i \\ q_i &\leqslant q_{max} \\ \mathbf{q}.\boldsymbol{e} &= 0 \\ \mathbf{M}\boldsymbol{q}.\boldsymbol{q} &\leqslant TE^2. \end{aligned} \tag{9}$$

5.2.1 A short digression on two limiting cases

This short section is devoted to the analysis of two limiting cases - not specific to the risk model under consideration - that highlight the role played by the constraints. This analysis leads to two approximate solutions to Problem (9) that are quite close to $\mathbf{A}^+$ and $\boldsymbol{\mu}^+$, respectively the positive part of the theoretical long-short portfolio and of the expected returns.

We now describe the two limiting cases just evoked.

Large TE bound When the tracking error constraint is large compared to the maximum weight constraint - Problem (9) approaches the limit problem

$$\begin{aligned} \sup \quad & \mathbf{q}.\mathbf{MA} \\ \boldsymbol{q} &= \boldsymbol{w} - \boldsymbol{I} \\ 0 &\leqslant w_i \\ q_i &\leqslant q_{max} \\ \mathbf{q}.\boldsymbol{e} &= 0. \end{aligned} \tag{10}$$

the solution $\mathbf{W}_{Large\,TE}$ of which is an explicit function of $\boldsymbol{\mu}^+ = (\mathbf{MA})^+$. Although the solution to (10) is not exactly $\boldsymbol{\mu}^+$, we will use $\boldsymbol{\mu}^+$ as a shorthand notation.

Large maximum weight constraint When the pointwise upper bound constraint is large compared to the TE bound, Problem (9) resembles

$$\begin{aligned} \sup & \quad \mathbf{q}.\mathbf{MA} \\ \boldsymbol{q} &= \boldsymbol{w} - \boldsymbol{I} \\ 0 &\leqslant w_i \\ \mathbf{q}.\boldsymbol{e} &= 0 \\ \mathbf{M}\boldsymbol{q}.\boldsymbol{q} &\leqslant TE^2. \end{aligned} \tag{11}$$

Due to non-locality, there is no explicit solution to (11), but a rather good approximate solution of (11) can be obtained by assuming that the set of positive portfolio weights coincides with those in $\mathbf{A}$, in which case the corresponding weight vector $\mathbf{W}_{No\,Max\,Weight}$ is well approximated by the long leg of the multi-score portfolio $\mathbf{A}^+$.

A conclusion to this analysis is that the solution $\mathbf{W}$ to (9) behaves like a mixture of the single factor portfolios and the two long-only portfolios $\boldsymbol{\mu}^+$ and $\mathbf{A}^+$. These two portfolios represent interaction terms formed by nonlinear combinations of the single factor portfolios.

6 Analyzing the cross-sectional risk model

In this section, we provide an empirical analysis, and provide some applications, of the cross-sectional risk model. Specifically, we analyse a multi-factor investment strategy in the S & P 500 universe. As already mentioned, the four style factors we consider are Low volatility, Momentum, Quality and Value. The model uses five macro-sector indicator portfolios, so that there are a total number of nine variables in the cross-sectional regression of the stock return alphas.

6.1 The cross-sectional returns

We study here the cross-sectional returns of the single factor portfolios. As already mentionned in Section 3.3, they are quite naturally interpreted as the realized returns of modified portfolios.

Comparing the cross-sectional and realized factor returns Figure 2 clearly shows that for each factor, both profiles look quite similar - the correlations between the realized and cross-sectional portfolio returns are all above 80%. Now, there are some differences, as discussed in detail in [1]. These differences come from the cross-sectional correlation of the single factor portfolios, that is, the fact that the Gram matrix $\mathbf{G}$ introduced in Section 3.3 is not diagonal. This is the only feature preventing the identification, up to normalization, of the realized and cross-sectional portfolio returns.

Factor returns correlation A more subtle difference, but one with major implications, lies in the decorrelation between the factor cross-

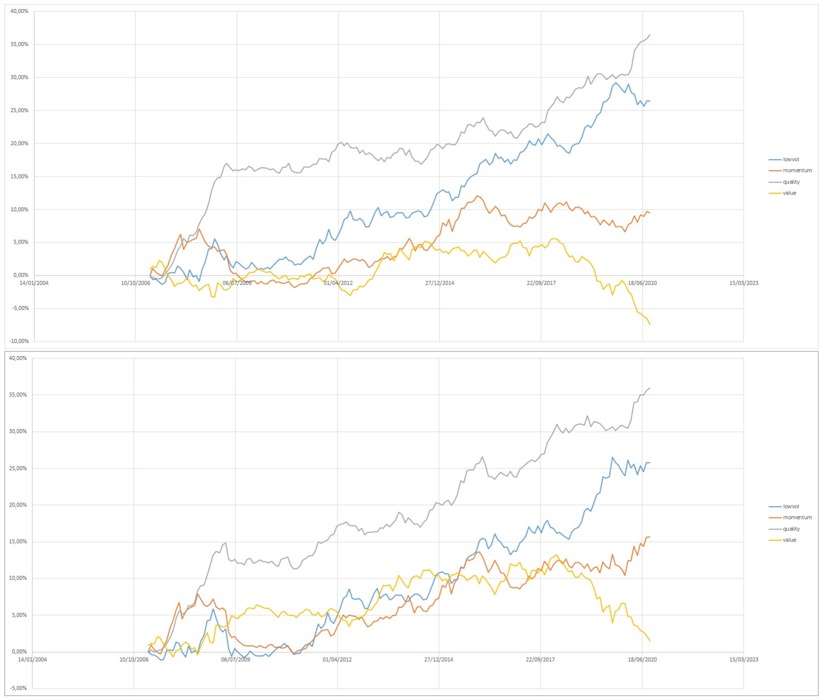

Figure 2: Realized and cross-sectional returns

sectional returns already evoked in Section 3.3. In fact, the best correlation model should rest on factor returns that would be as little correlated as possible, so that the correlation structure between stocks can be fully explained by the scores, see Formula (6).

The global correlation matrix shown in Figure 3 indicates that style factor are "reasonably" decorrelated[5], and the correlations are forced accordingly to 0 in the model. However, the sector-sector correlations may be high, as are some factor-sector correlations. In the risk model, it seems necessary to keep these sector-sector and factor-sector correlations to their historical level in order to generate a realistic covariance matrix. This is consistent with the observations made in Section 4.2 about the relevance of sectors to explain the correlation between stocks.

A time-varying analysis of these correlations, comparing them with the realized long-short portfolio returns correlations, is another argument in favour of the cross-sectional risk model:

Figure 4 clearly shows that the absolute correlation level between cross-sectional factor returns is generally lower than that between the corresponding long-short portfolio realized returns.

[5]The 95% confidence interval for a 0 correlation coefficient with this sample size is approximately 15%, see e.g. [6]

	lowvol	momentum	quality	value	CYCLICAL	DEFENSIVE	ENERGY & INDUSTRIALS	FINANCIALS	IT & TELECOM
lowvol	100%	20%	7%	-13%	-33%	22%	-56%	-12%	-43%
momentum	20%	100%	8%	-22%	-54%	21%	-19%	-51%	-4%
quality	7%	8%	100%	-35%	-18%	-9%	-39%	-38%	2%
value	-13%	-22%	-35%	100%	23%	7%	24%	40%	-11%
CYCLICAL	-33%	-54%	-18%	23%	100%	-17%	38%	47%	27%
DEFENSIVE	22%	21%	-9%	7%	-17%	100%	-21%	-14%	-14%
ENERGY & INDUSTRIALS	-56%	-19%	-39%	24%	38%	-21%	100%	8%	19%
FINANCIALS	-12%	-51%	-38%	40%	47%	-14%	8%	100%	-5%
IT & TELECOM	-43%	-4%	2%	-11%	27%	-14%	19%	-5%	100%

Figure 3: The cross-sectional returns correlation matrix

Figure 4: Sliding correlations

6.2 Analyzing the covariance matrix

The covariance matrix is the most useful and important tool in the portfolio optimization process, and one of our main goals is to generate a covariance that not only incorporates explicitly the scores of each stocks in the covariance structure, but also eases the optimization process by exhibiting some good conditioning properties.

Spectrum of the covariance matrix The properties of matrix $\mathbf{M}$ in Equation (6), especially those of its spectrum, are very important. Thanks to its rather parsimonious construction and the importance of its diagonal (idiosyncratic) part, one can expect some good conditioning properties. The results in this section show that the matrix actually behaves in a very satisfactory way. In particular, the lower end of its spectrum is 10 to 100 times higher that that of the empirical variance-covariance matrix, a fact that greatly reduces the usual small eigenvalue problems that occur in portfolio optimization.

Figure 5 shows the ratio between eigenvalues of same rank in the empirical and model-based covariance matrices. It clearly demonstrates that the condition number of the cross-sectional model-based covariance matrix is much better than that of the empirical covariance matrix.

A more precise analysis actually confirms that the small eigenvalues from Equation (6) are much larger (2 to 3 orders of magnitude) than the empirical ones and therefore, that inverting the covariance matrix becomes a much safer task with the cross-sectional model.

6.3 The theoretical long-short portfolio

In this section, we compare the performances of four multi-factor long-short candidate portfolios. As explained in Section 6, they are obtained *via* equal risk budgets or contributions from the original or the cross-sectional long-short portfolios. Their cumulative performances are shown on Figure 6.

Clearly, the four portfolio performances are extremely similar, and the apparently better performance of the "LS ERB" portfolio - equal risk budgeting on the original single factor portfolios - is only due to its higher *ex post* volatility. All four multi-factor portfolios actually have very similar information ratios all approximately equal to 1. Therefore, one needs to use other criteria to discriminate between the four candidates.

6.4 The long-only investment portfolio

In this final section, we provide some examples of long-only investment portfolios built using the cross-sectional, factor-based risk model to solve Problem (9). The graphs show the cumulative performances of the long-only and other related portfolios: the theoretical long-short portfolio and the two additional portfolios introduced in Section 5.2.1.

One of our goals here is to exemplify the role played by the tracking error (TE) constraint: for a TE set to a 5% annual volatility level, one obtains the investment portfolio shown on Figure 7, which is very similar to the $\boldsymbol{\mu}^+$ portfolio introduced in Equation (10). In fact, a direct inspection of the optimal portfolio active weights shows that the maximum weight constraint is very often saturated, leading to a correlation between the long-only portfolio and the $\boldsymbol{\mu}^+$ portfolio of 88%.

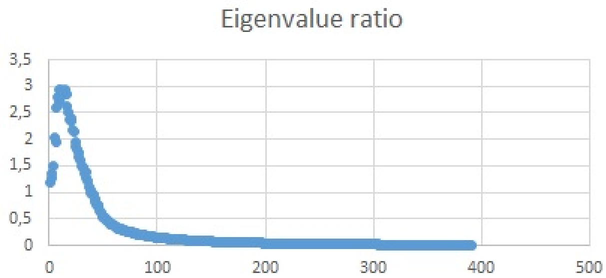

Figure 5: Comparing the empirical and cross-sectional covariance spectra

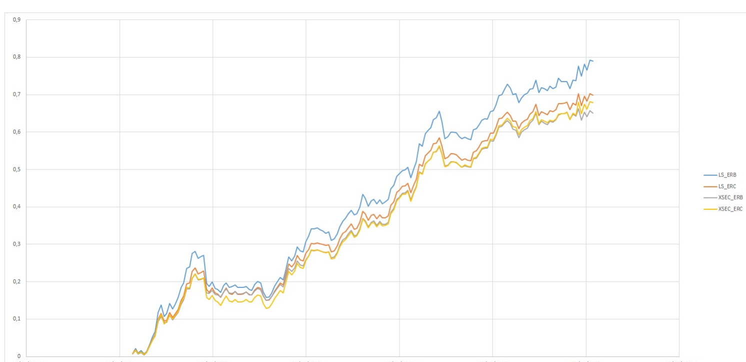

Figure 6: The cumulative performances of the multi-factor long-short portfolio candidates

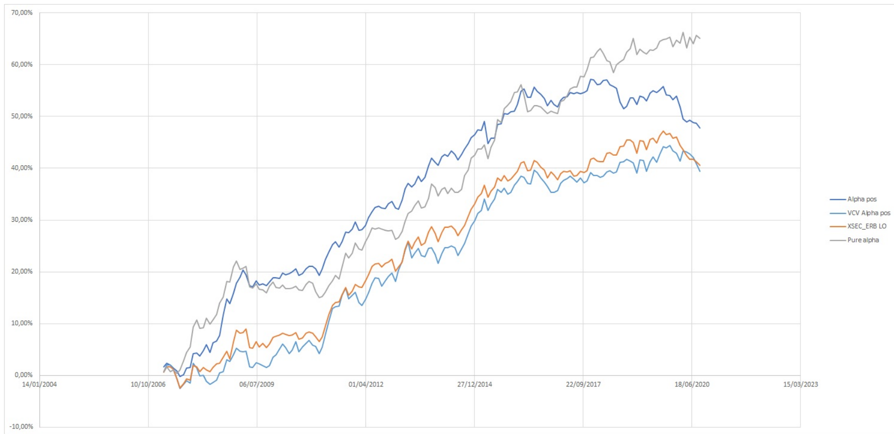

Figure 7: Investment portfolio with high tracking error

On the other hand, setting the tracking error at the more reasonable level of 3.5% yields the more mixed results shown on Figure 8. As the number of weights saturating the maximum pointwise constraint is much lower, the correlation between the long-only portfolio and the $\mathbf{A}^+$ portfolio now becomes very high (92%).

Figure 8: Investment portfolio with moderate tracking error

References

[1] F. Abergel and T. Heckel. Performance attribution for multi-factorial equity portfolios. *submitted*, 2021.

[2] Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch. Paschalidis. Inverse optimization: A new perspective on the black-litterman model. *Operations Research*, 60(6):1389–1403, 2012.

[3] B. Bruder and T. Roncalli. Managing risk exposures using the risk budgeting approach. 2012.

[4] K. Daniel and S. Titman. Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, 52(1):1–33, 1997.

[5] E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 3, 1993.

[6] R. A. Fisher. On the "probable error' of a coefficient of correlation deduced from a small sample. *Metron*, 3, 1921.

[7] M. Haug, G. Iyengar, and I. Song. A generalized risk budgeting approach to portfolio construction. *Journal of Computational Finance*, 21, 2017.

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

[9] G. He and R. Litterman. The intuition behind black-litterman model portfolios. 2002.

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

[11] R. Litterman and Goldman Sachs Quantitative resources group. *Modern Investment Management: An Equilibrium Approach*. Wiley finance, 2003.

[12] S. Maillard, T. Roncalli, and J Teiletche. On the properties of equally-weighted risk contributions portfolios. 2009.

[13] A. Saxena and R. A. Stubbs. The alpha alignment factor: a solution to the underestimation of risk for optimized active portfolios. *Journal of Risk*, 15, 2013.

List of Figures