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

Frédéric Abergel
Quantitative research group
BNP Paribas Asset Management

BSDE 2022 conference
Industrial afternoon session

28 june 2022

Plan I

1. Introduction: Factor investing in a nutshell

2. Factor-based risk model: General methodology
   - Factor model for a stock's alpha
   - The cross-sectional factor portfolios
   - Variance-covariance structure
   - Analyzing the VCV matrix

3. Portfolio construction
   - Building a long-only, investable portfolio

4. Applications and examples
   - The single factor portfolios
   - The long-only portfolios

5. Concluding remarks

Factor investing in a nutshell

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) asserts that the performance of a stock should be determined by a single stock characteristic, the **beta**.

However, it has long been accepted, see e.g. [Fama and French, 1993] [Jegadeesh and Titman, 1993], that other factors known as Value, Low Volatility, Quality and Momentum play a role in explaining stock returns.

Factors in a nutshell

- **Value** characteristics such as the price-to-book or the price-to-earnings of a company measure the relative cheapness of a stock. On average, over time, cheaper stocks tend to out-perform other stocks
- Stocks with a low **Volatility** generate at least comparable returns to riskier stocks, rendering them more attractive: same returns in the medium to long-term with less uncertainty
- **Quality** stocks, e.g. the most profitable companies, tend to generate higher returns than other stocks
- Stocks with the strongest **Momentum**, 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

Factor investing in a nutshell

Single factor portfolios are theoretical portfolios built from suitable ranking, aggregation, normalization and neutralization. These long-short portfolios have weights that reflect the score of an asset according to one particular member of the factor family.
Multi-factor investing consists in building, based on these single factor portfolios, an investment strategy.
In the asset management world, this process can be decomposed into two major steps:

- The generation of a global score for each stock, leading to a theoretical multi-factorial long-short portfolio
- The construction of an investable portfolio meeting a set of constraints: long-only, controlled volatility, diversification, controlled transaction costs...

Both steps require the use of a **risk model** to perform suitable asset allocations.

Factor investing in a nutshell

A **risk model** consists in the design of the dependency structure between assets.
Starting from the CAPM model

$$R_i(t) = \beta_i RM(t) + \alpha_i$$

one seeks a model for the dependency structure of the stock return alphas α_i. In a Gaussian framework, this is simply their covariance matrix.
This presentation addresses the design of this dependency structure in the context of multi-factor investing. It is based on the recently published paper [Abergel et al., 2022]

Plan I

Plan I

1. Introduction: Factor investing in a nutshell
2. Factor-based risk model: General methodology
   - Factor model for a stock's alpha
   - The cross-sectional factor portfolios
   - Variance-covariance structure
   - Analyzing the VCV matrix
3. Portfolio construction
   - Building a long-only, investable portfolio
4. Applications and examples
   - The single factor portfolios
   - The long-only portfolios
5. Concluding remarks

Factor model for a stock's alpha

The assets' alphas are explained *via* a cross-sectional regression of the returns against a small number of explanatory variables:

- Long-short factor portfolio weights
- Industrial sector indicatrixes

In a continuous-time framework, the model would be written as:

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

In vector notations:

$$d\boldsymbol{X}(t) = \sum_{s=1}^{S} \mathbf{1}_s(t) d\mu_s(t) + \sum_{k=1}^{K} \mathbf{z}_k(t) d\lambda_k(t) + d\boldsymbol{\epsilon}(t) \quad (2.1)$$

Factor model for a stock's alpha

We switch to the discrete-time version of (2.1) for monthly returns:

$$\boldsymbol{X} = \sum_{s=1}^{S} \mathbf{1}_s \mu_s + \sum_{k=1}^{K} \mathbf{z}_k \lambda_k + \boldsymbol{\epsilon}$$

In a more compact form:

$$\boldsymbol{X} = \sum_{k=1}^{K+S} \mathbf{z}_k \lambda_k + \boldsymbol{\epsilon} \tag{2.2}$$

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

Practical implementation

For modularity reasons, and in order to better distinguish between factorial and other effects, we actually perform a two-step regression

- first, an OLS regression of the alphas against the factor portfolios
- then, an OLS regression of the residuals against the macro-sector indicatrix portfolios

That way, one can add some extra risk factors, e.g. region indicatrixes, finer-grain sectors, supply chain clusters... without changing the factor-based part of the model.
It also simplifies the correspondence between the basic long-short factor portfolios and the **cross-sectional** factor portfolios that we now introduce.

Plan I

The cross-sectional factor portfolios

The building blocks of factor investing are the **single factor** portfolios, and the cross-sectional approach paves the way to a new understanding of single factor portfolios.

Consider Equation (2.2) and introduce the Gram matrix **G** of the factor portfolios

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

The solution to the OLS regression is

$$\mathbf{\Lambda} = \mathbf{G}^{-1}\mathbf{RLS} \tag{2.3}$$

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

The cross-sectional factor portfolios

The λ_k are the **cross-sectional** returns of the original factor portfolios. They can also be interpreted as the **realized** returns of modified factor portfolios

$$\mathbf{y}_k = \mathbf{z}_k{}^t(\mathbf{G}^{-1}) \equiv \mathbf{z}_k \mathbf{G}^{-1}$$

We call these new portfolios $\mathbf{y}_k$ the **cross-sectional factor portfolios**.

We shall see that their returns are very weakly correlated over time, as opposed to those of the original long-short portfolios, so that they provide natural candidates for factor investing in independent (uncorrelated) factors.

Plan I

Variance-covariance structure

The model VCV matrix coming from Equation (2.2) has two main components:

- A term coming from the variance-covariance matrix of the cross-sectional return $\lambda_k(t)$
- An idiosyncratic term given by the variance of $\epsilon_i(t)$

More specifically:

- The factor- and sector-based alpha 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 **M** is defined by

$$\mathbf{M}_{ij} = \mathbf{F}_{ij} + \mathbf{D}_{ij}. \tag{2.4}$$

One of the main features of this model is the alignment between the VCV matrix and the assets' scores, see e.g. [Saxena and Stubbs, 2013] for a related analysis.

Plan I

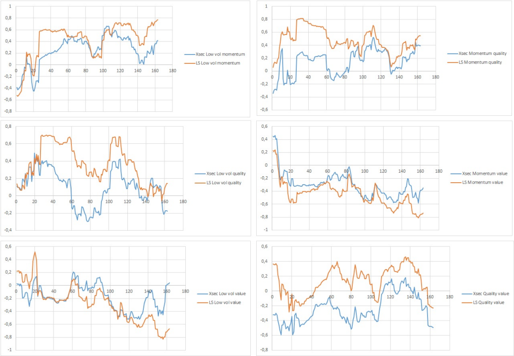
Factor returns correlation

The best possible cross-sectional model should yield factor returns that are as little correlated as possible. The full period correlation matrix indicates that the cross-sectional factor portfolios are "reasonably" decorrelated over time. Note however that some sector-sector and factor-sector correlations remain significant.

	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%

Factor returns correlation

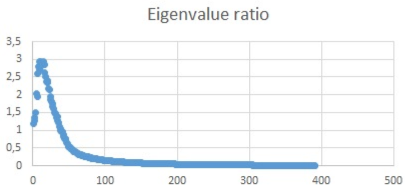
A time-dependent analysis provides more arguments in favour of the cross-sectional portfolios (or: the cross-sectional returns of the single factor portfolios)



Variance-covariance spectrum

The VCV matrix **M** in Equation (2.4) is used extensively during portfolio construction, and the properties of its spectrum are therefore very important.

The graph below shows the ratio of eigenvalues of same rank between the empirical variance-covariance matrix and **M**. A more precise analysis confirms that inverting the VCV matrix becomes much safer - shrinkage is no more necessary.



Plan I

Plan I

Building a long-only, investable portfolio

The optimization program uses the model-based VCV matrix and a set of expected returns built from *reverse optimization*, see e.g. [Litterman and Quantitative resources group, 2003].
Denoting by $\boldsymbol{w}$ the vector of portfolio weights and $\mathbf{q}$ that of active weights $\mathbf{q} \equiv \mathbf{w} - \mathbf{I}$, the general program is the following:

$$
\begin{aligned}
\sup \quad & \mathbf{q}.\mathbf{M}\alpha \\
0 \quad & \leqslant \quad w_i \\
q_i \quad & \leqslant \quad c \\
\mathbf{q}.\boldsymbol{e} \quad & = \quad 0 \\
\mathbf{M}\boldsymbol{q}.\boldsymbol{q} \quad & \leqslant \quad TE^2
\end{aligned}
\tag{3.1}
$$

and it can be useful to try and analyze some simplified problems highlighting the role played by the constraints.

Large TE bound

When the tracking error bound is large compared to the maximum weight constraint, Problem(3.1) approaches the limit problem

$$\begin{aligned} \sup & \quad \mathbf{q}.\mathbf{M}\alpha \\ 0 & \leqslant w_i \\ q_i & \leqslant q_{max} \\ \mathbf{q}.\boldsymbol{e} & = 0 \end{aligned} \tag{3.2}$$

the solution of which can be written as an explicit function of the expected returns $\mathbf{M}\alpha$: Assuming for simplicity that $q_{max} = \frac{1}{NMAX}$, $NMAX$ an integer, then the solution is obtained by assigning the weight c_{max} to the $NMAX$ stocks taking the largest values in the expected returns $\mathbf{M}\alpha$.

No maximum weight constraint

When no pointwise upper bound constraint is enforced, Problem (3.1) becomes

$$
\begin{aligned}
\sup \quad & \mathbf{q}.\mathbf{M}\alpha \\
0 \leqslant \; & w_i \\
\mathbf{q}.\boldsymbol{e} = \; & 0 \\
\mathbf{M}\boldsymbol{q}.\boldsymbol{q} \leqslant \; & TE^2
\end{aligned}
\tag{3.3}
$$

Due to non-locality (**M** is not diagonal) there is no explicit solution to (3.1), but the problem can be analyzed as a variational inequality:

$$
\begin{aligned}
(\mathbf{Mq})_i - c_1(\mathbf{M}\alpha)_i - c_2 &= 0, \; w_i > 0 \\
(\mathbf{Mq})_i - c_1(\mathbf{M}\alpha)_i - c_2 &\leqslant 0, \; w_i = 0
\end{aligned}
$$

An approximate solution is obtained by assuming that the set of positive active weights is the same as the set of positive weights in the theoretical multi-factorial portfolio.

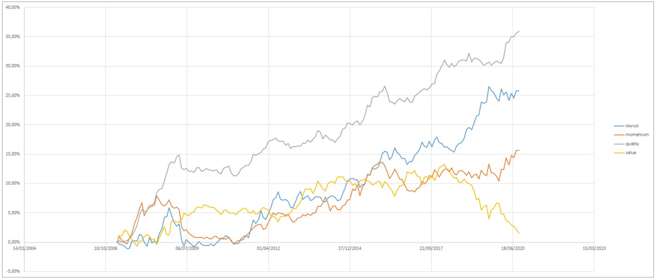
Plan I

Some numerical results for the S&P500 universe

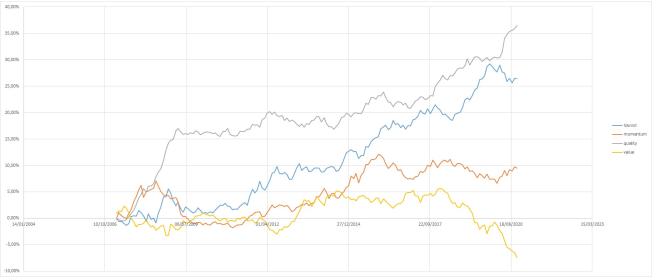
We specialize the cross-sectional risk model to the S & P 500 universe.
There are four style factors "Low volatility", "Momentum", "Quality" and "Value", and five macro-sectors "Cyclical", "Defensive", "Energy and industrials", Financials, "IT and telecom", giving a total number of nine factors for the cross-sectional regression

Plan I

The original factor portfolios



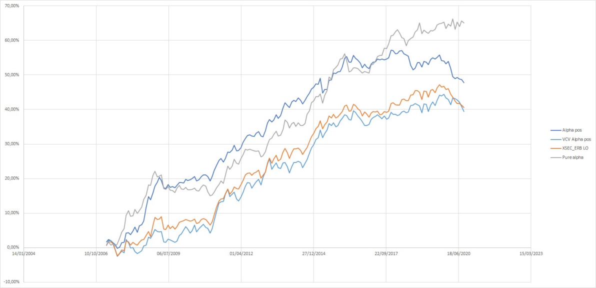
The cross-sectional factor portfolios



Plan I

Focusing on the long-only portfolio

For a TE level set to 5%, one obtains the following investment portfolio, very close to the first approximate solution introduced in Section 1



A direct inspection of the weights shows that the maximum weight constraint is very often saturated, leading to a correlation between the long-only portfolio and the approximate portfolio of 88%.

On the other hand, setting the TE level at the more reasonable level of 3.5% yields different results:

Now, the correlation between the long-only portfolio and the second approximate solution introduced in Section 1 is very high (92%), and the number of weights saturating the maximum pointwise constraint is much lower.

Plan I

Summary and concluding remarks

- A cross-sectional risk model has been built from style factors and sectors
- The model is built from cross-sectional factor returns that exhibit good decorrelation properties
- The covariance of two stocks in the model is an explicit function of the stocks' scores, thereby making the model well aligned with the factors
- The model-based VCV matrix has much better spectral properties that the empirical stock VCV matrix
- Analyzing the long-only portfolio construction highlights three target portfolios

References I

Abergel, F., Bellone, B., and Soupé, F. (2022).
A factor-based risk model for multifactor investment strategies.
Journal of Risk, 24.

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

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

Litterman, R. and Quantitative resources group, G. S. (2003).
Modern Investment Management: An Equilibrium Approach.
Wiley finance.

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