Climate Risk Hedging

Thomas Lorans

May 31, 2024

Contents

Introduction

Chapter 1

Factor Mimicking Portfolios

1.1 Risk is Innovation

We have a vector of K factors of risks F_{t+h}, with h the forecast horizon. Investors form expectations of these factors at time $t-1$ and adjust their expectations at time t based on new information. The change in expectations is given by:

$$\tilde{F}_{t+h} = F_{t+h} - \phi \tag{1.1}$$

where $\tilde{F}_{t+h}$ is the *innovation* in the factors of risks.

1.2 Innovation and Unexpected Returns

On the other hand, we have the *unexpected* returns $\tilde{R}_t$:

$$\tilde{R}_t = R_t - \mu \tag{1.2}$$

The main assumption behind factor mimicking portfolios is that the innovation $\tilde{F}_{t+h}$ is reflected in the unexpected returns $\tilde{R}_t$:

$$\tilde{R}_t = B\tilde{F}_{t+h} + \varepsilon_t \tag{1.3}$$

where B is a $N \times K$ matrix of factor loadings, ε_t is a $N \times 1$ vector of mean zero disturbances.

It means that investors reprice assets (unexpected returns $\tilde{R}_t$) based on the arrival of new information on the factors of risks (innovation $\tilde{F}_{t+h}$).

1.3 Linear Factor Model

If:

$$R_t = \mu + \tilde{R}_t \tag{1.4}$$

Then, substituting $\tilde{R}_t$, we have the following factor model:

$$R_t = \mu + B\tilde{F}_{t+h} + \varepsilon_t \tag{1.5}$$

with R_t a $N \times 1$ vector of asset returns, μ a $N \times 1$ vector of expected returns, B a $N \times K$ matrix of factor loadings, F_{t+h} a $K \times 1$ vector of factor innovations and ε_t a $N \times 1$ vector of mean zero disturbances.

1.4 Factor Mimicking

The vector of weights w_k is the solution to the following optimization problem:

$$\begin{aligned} \min_{w_k} \quad & \frac{1}{2} w_k^T \Sigma w_k \\ \text{subject to} \quad & B^T w_k = \beta_k \end{aligned} \tag{1.6}$$

where B is the $N \times K$ matrix of factor loadings, β_k is the $K \times 1$ vector of factor exposures, with the k-th element equal to 1 and the other elements equal to $\beta_{k,l}$, and Σ is the $N \times N$ covariance matrix of asset returns.

We can form the Lagrangian:

$$\mathcal{L}(w_k, \lambda) = \frac{1}{2} w_k^T \Sigma w_k - \lambda_k^T (B^T w_k - \beta_k) \tag{1.7}$$

where λ_k is the $K \times 1$ vector of Lagrange multipliers.

The first order condition is:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w_k} &= \Sigma w_k - B\lambda = 0 \\ \Rightarrow w_k &= \Sigma^{-1} B \lambda_k \end{aligned} \tag{1.8}$$

Substituting w_k in the constraint, we have:

$$\begin{aligned} B^T w_k &= \beta_k \\ B^T \Sigma^{-1} B \lambda_k &= \beta_k \\ \Rightarrow \lambda_k &= (B^T \Sigma^{-1} B)^{-1} \beta_k \end{aligned} \tag{1.9}$$

Substituting λ_k in w_k, we finally have the solution to the optimization problem:

$$w_k^* = \Sigma^{-1} B (B^T \Sigma^{-1} B)^{-1} \beta_k \tag{1.10}$$

1.5 Risk Premia

1.6 Conclusion

In what follows, we will focus on the case of climate risk factors. First stage is to identify how to measure climate risk factors.

Chapter 2

Measuring Climate Risk Factors

A key challenge in implementing a dynamic hedging strategy for climate risk is to construct a time series that captures news about long-term climate risk.

We can start from the observation that when there are events that plausibly contains information about changes in climate risk, this will likely leads to newspaper coverage of these events. Newspapers may even be the direct source that investors use to update their beliefs about climate risk.

2.1 Representing Text as Data

Text is high dimensional. Suppose we have a bunch of documents, each of which is w words long. Each word is drawn from a vocabulary of p possible words. The unique representation of these documents has dimension p^w.

Analysis can summarized in three steps:

1. Represent raw text D as a numerical array C
2. Map C to predicted values $\hat{V}$ of unknown outcomes V
3. Use $\hat{V}$ in subsequent analysis

The first step in constructing C is to divide the raw text D into individual documents D_i. The way to divide the raw text is dictated by the value of interest V. If V is daily stock price, it might makes sense to divide the raw text into daily news articles.

To begin with the transformation from raw text D to a numerical array C, we can first count the number of times each word appears in each document,

$c_{i,j}$. It results into a matrix C of size $n \times p$ where n is the number of documents and p is the number of unique words in the vocabulary. Each row of C refers to a document i, and each column refers to a word j.

Example X.1.

D_1	a rose is still a rose
D_2	there is no there there
D_3	rose is a rose is a rose is a rose

Table 2.1: Examples of individual documents D_i

$i \setminus j$	a	rose	is	still	there	no
1	2	2	1	1	0	0
2	0	0	1	0	3	1
3	3	4	3	0	0	0

Table 2.2: Term frequency matrix C

2.2 Dictionary-based Mapping

Dictionary-based methods are used to map the counts c_i to outcomes v_i. It specify $\hat{v}_i = f(c_i)$ where f is a function pre-specified. Dictionary-based methods heavily rely on prior information about the function mapping c_i to outcomes v_i. They are more appropriate when prior information is strong and reliable and where information in the data is weak. An example is a case where the outcomes v_i are not observed for any i, so there is no training data available.

Example X.1. Suppose we have a dictionary-based method that maps the counts c_i of to outcomes v_i. The dictionary is a list of words for a specific category.

Category	Dictionary
Positive	good, great, excellent
Negative	bad, terrible, awful

Table 2.3: Example of a dictionary-based method

We have the following documents D_i:

D_1	good is great
D_2	bad is terrible
D_3	good is bad

Table 2.4: Example of documents D_i

The matrix C is:

$i \setminus j$	good	great	bad	terrible	is
1	1	1	0	0	1
2	0	0	1	1	1
3	1	0	1	0	1

Table 2.5: Term frequency matrix C

Mapped to the dictionary, it becomes:

$i \setminus k$	Positive	Negative
1	2	0
2	0	2
3	1	1

Table 2.6: Mapped matrix C

We define the function $f(c_i)$ as:

$$f(c_i) = \begin{cases} \text{Positive} & \text{if } \sum_j c_{i,j} \in \text{Positive} \\ \text{Negative} & \text{if } \sum_j c_{i,j} \in \text{Negative} \\ \text{Neutral} & \text{otherwise} \end{cases} \tag{2.1}$$

Chapter 3

Climate Risk Mimicking Porfolios

3.1 Two-Pass Fama-MacBeth

3.2 Maximum Correlation Portfolio

Chapter 4

Conclusion

More generally can be applied to other ESG risks. See biodiversity risk from Giglio et al.