Climate Risk Hedging

Thomas Lorans

May 27, 2024

Contents

Introduction **v**

1 Hedging Climate Risk with Climate News **1**
1.1 Climate Long-Run Risk 1
1.2 Uncovering Long-Run Climate Risk with Changes in Expectation 2
1.3 Climate Factor Signal 5
1.4 Further Reading 5

2 Measuring Climate News **7**
2.1 Representing Text as Data 7
2.2 Dictionary-based Mapping 8
2.3 Further Reading 10

3 Efficient Mimicking Porfolio **11**
3.1 Portfolio Optimization 11
3.2 Climate Efficient Mimicking Portfolio 13
3.3 Further Reading 13

4 Conclusion **15**

Introduction

Chapter 1

Hedging Climate Risk with Climate News

Because of the long-term nature of climate risk, standard futures or insurance contracts in which one party agrees to pay the other in the event of a climate disaster are difficult to implement. Rather than buying a security that pays off in the event of a climate disaster, you can construct portfolio whose short-term returns hedge *news* about climate risk. By hedging, period by period, the innovations in news about long-run climate risk, an investor can ultimately hedge her long-run exposure to climate risk.

1.1 Climate Long-Run Risk

Climate risk is a long-run risk. It is usually separated into two components:

- **Physical risk**: the risk that climate change will have direct effects on the value of assets. For example, a rise in sea levels could affect the value of real estate.

- **Transition risk**: the risk that the transition to a low-carbon economy will affect the value of assets. For example, a carbon tax could affect the value of fossil fuel companies.

While global warming already has physical effects, most of the potential damages are still in the future. The transition to a low-carbon economy is also a long-run process, with a lot of uncertainty about the timing and the nature of the transition.

CLIMATE SCENARIOS.
$CC_{t+\tau}$ is the climate risk at time $t+\tau$.

1.2 Uncovering Long-Run Climate Risk with Changes in Expectation

We are facing a strong uncertainty about $CC_{t+\tau}$, the climate risk at time $t+\tau$, and it is not directly observable.

We can draw inspiration from the literature on nowcasting and forecasting to estimate the current expectation of $CC_{t+\tau}$, with the idea of *factor mimicking*, following Lamont (2001) [?].

The first core idea is to replace some variables with a linear combination of other variables. More specifically, **some variable of interest can be written as a portfolio of *tradable assets***. It can be used to proxy economic variable that are not directly observable with tradable assets. We can construct, from financial assets, a "matching portfolio" of some economic factor that is not directly observable.

The second core idea is that we can **uncover expectations about any variable correlated with future cash flows and discount rates by looking at the unexpected returns of tradable assets.**

For climate risk, let's start by defining the current expectation of $CC_{t+\tau}$ as:

$$E_t(CC_{t+\tau}) = E_{t-1}(CC_{t+\tau}) + \Delta E_t(CC_{t+\tau}) \tag{1.1}$$

That is, the current expectation is the previous expectation plus the news or "surprise" about the climate risk. On a similar vein, we can define realized returns as:

$$R_t = E_{t-1}(R_t) + \tilde{R}_t \tag{1.2}$$

with $\tilde{R}_t$ the unexpected returns. It's simply the difference between the expected returns and the actual returns $\tilde{R}_t = R_t - E_{t-1}(R_t)$.

Our key assumption now is that **innovations in returns (unexpected returns) reflect innovations in expectations about the climate risk**, such that:

$$\Delta E_t(CC_{t+\tau}) = \beta^T \tilde{R}_t + \epsilon_t \tag{1.3}$$

If the climate risk $CC_{t+\tau}$ is correlated with future cash flows and discount rates, then we may find something in the β, relating news reflected in the unexpected returns. This is based on the assumption that the unexpected returns reflect news about the future cash flows and discount rates (*i.e.* about $\Delta E_t(CC_{t+\tau})$).

The other key assumption is that the long-run climate risk is simply the sum of the news about the climate risk:

$$E_t(CC_{t+\tau}) = \sum_{k=1}^{\tau} \Delta E_t(CC_{t+k}) \tag{1.4}$$

which is true if $E_t(CC_{t+\tau}) = E_{t-1}(CC_{t+\tau}) + \Delta E_t(CC_{t+\tau})$.

Example X.1. Say you want to estimate current (this is *nowcasting*) GDP or inflation. You can construct the portfolio of assets that best mimics the movements of GDP or inflation. Once you've run your regression, you can use your estimate of returns to predict the macro variable. Let's say we don't have individual stock returns and we want to estimate the market return. All what we have is the return of K industry portfolios. We can estimate the market return as a linear combination of the industry returns:

$$R_{m,t} = \beta_{\text{energy}} R_{\text{energy},t} + \beta_{\text{financials}} R_{\text{financials},t} + \ldots + \beta_k R_{k,t} + \epsilon_t \tag{1.5}$$

It looks very much like a portfolio, with the estimated β as the weights of the assets.

Now, let's say we want to form a portfolio that mimics the behavior of a *factor signal* y. Specifically, our target is "news" or *innovations* about the signal, defined as the difference between the current expectation and the previous expectation:

$$\Delta E_t(y_{t+1}) := E_t(y_{t+1}) - E_{t-1}(y_{t+1}) \tag{1.6}$$

In can be for example the news that the market learns about the industrial production in May about the industrial production in June.

The mimicking portfolio portfolio returns are:

$$R_{y,t} = w^T R_t \tag{1.7}$$

where R_t is the vector of returns of the tradable assets. You construct the mimicking portfolio with *unexpected returns* of the tradable assets. Unexpected returns are actual returns minus expected returns:

$$\tilde{R}_t := R_t - E_{t-1}(R_t) \tag{1.8}$$

with the assumption that **the expected returns are a linear function of factors F_t**:

$$E_{t-1}(R_t) = \gamma^T F_t \tag{1.9}$$

The porfolio weights of the mimicking portfolio of y are chosen so that $\tilde{R}_{y,t}$ is as close as possible to $\Delta E_t(y_{t+1})$ (maximally correlated).

To do this, the key assumption is that **innovations in returns (unexpected returns) reflect innovations in expectations about the factor signal**, such that:

$$\Delta E_t(y_{t+1}) = \beta^T \tilde{R}_t + \epsilon_t \tag{1.10}$$

If the factor signal y is correlated with future cash flows and discount rates, then we may find something in the β, relating news reflected in the unexpected returns. Again, this is based on the assumption that the unexpected returns reflect news about the future cash flows and discount rates (*i.e.* about $\Delta E_t(y_{t+1})$).

Recalling that the returns are $R_t = E_{t-1}(R_t) + \tilde{R}$, we can therefore rewrite it with the factors:

$$R_t = \gamma^T F_t + \tilde{R} \tag{1.11}$$

and then includes the innovations in the factor signal:

$$R_t = \gamma^T F_t + \eta \Delta E_t(y_{t+1}) + u_t \tag{1.12}$$

What we have here? **The returns of any asset can be written as a function of its expected returns ($\gamma^T F_t$) and the unexpected returns. The unexpected returns are decomposed into the news about the factor signal $\Delta E_t(y_{t+1})$ and uncorrelated errors (u_t).**

1.3 Climate Factor Signal

In the case of climate risk, we have:

$$\Delta E_t(CC_{t+k}) = E_t(CC_{t+k}) - E_{t-1}(CC_{t+k}) \tag{1.13}$$

with CC_{t+k} the climate risk at an undefined horizon k. We have seen in the chapter 1 how to use text to proxy for the market expectation of the climate risk $E_t(CC_{t+k})$.

Example X.1. *illustrates use of demean vs. AR(1) to estimate*

Therefore, we can estimate:

$$R_{i,t} = \beta_i \Delta E_t(CC_{t+k}) + \gamma_i^T \text{Factors}_t + \epsilon_{i,t} \tag{1.14}$$

The β_i for each portfolio is the signal upon which we want to derive our mimicking portfolio. The higher is β_i, the higher will be the weight of the portfolio in the mimicking portfolio.

However, if β_i is negative, it means that we should short the portfolio. In practice, it is less common to short portfolios, so we can set $\beta_i = 0$ if it is negative. In that case, the mimicking portfolio would be a *long-only* portfolio.

A simple method to go from the β_i to the weights of the mimicking portfolio is to normalize it:

$$w_i = \frac{\tilde{\beta}_i}{\sum_{i=1}^N \tilde{\beta}_i} \tag{1.15}$$

with $\tilde{\beta}_i = \max(\beta_i, 0)$.

The vector of weights w is already an intuitive long-only mimicking portfolio: it simply **weights the portfolios based on their positive β_i on the climate innovation signal $\Delta E_t(CC_{t+k})$ we wish to mimic**.

1.4 Further Reading

Tracking portfolio: Lamont (2001) [**?**]

Regression: Climate risk mimicking: Engle *et al.* (2020) [**?**] Alekseev *et al.* (2022) [**?**]

Chapter 2

Measuring Climate News

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:

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

2.3 Further Reading

Use of text: *Text as Data* by Gentzkow *et al.* (2019) [?] *Narrative Asset Pricing* by Bybee *et al.* (2023) [?]

Climate series: Engle *et al.* (2020) [?] Apel *et al.* (2023) [?]

Chapter 3

Efficient Mimicking Porfolio

We now want to improve upon this first portfolio by including more sophisticated optimized mimicking portfolio construction, by **taking into account the information on the covariance between the assets**. Indeed, even though the mimicking portfolio we have seen in the previous chapter is very intuitive, it can be **suboptimal depending on the risk appetite of the investor**.

3.1 Portfolio Optimization

Let's consider an universe of N assets. We have $w = \begin{bmatrix} w_1 & w_2 & \dots & w_N \end{bmatrix}^T$ the vector of weights in the portfolio. We suppose the portfolio is fully invested, i.e. $w^T\mathbf{1} = 1$. We have $R = \begin{bmatrix} R_1 & R_2 & \dots & R_N \end{bmatrix}^T$ the vector of returns of the assets.

The return of the portfolio is given by:

$$R_p = \sum_{i=1}^{N} w_i R_i = w^T R \tag{3.1}$$

The vector of expected asset returns is denoted by $\mu = E(R)$. The covariance matrix of the asset returns is denoted by Σ:

$$\Sigma = E[(R-\mu)(R-\mu)^T] \tag{3.2}$$

The variance of the portfolio is given by:

$$
\begin{aligned}
\sigma_p^2 &= E[(R_p - E(R_p))^2] \\
&= E[(w^T R - w^T \mu)^2] \\
&= E[(w^T (R - \mu))^2] \\
&= E[w^T (R - \mu)(R - \mu)^T w] \\
&= w^T E[(R - \mu)(R - \mu)^T] w \\
&= w^T \Sigma w
\end{aligned}
\tag{3.3}
$$

Example X.1. *illutrates to show why it is quadratic*

The problem of the investor can be formulated as:

1. **Maximize the expected return of the portfolio under a volatility constraint** ($\sigma_p \leq \sigma^*$)

2. **Minimize the volatility of the portfolio under a return constraint** ($\mu_p \geq R^*$)

The key idea of Markowitz (1956) is to combine the two objectives into a single quadratic optimization problem:

$$
\begin{aligned}
\min_{w} \quad & \frac{1}{2} w^T \Sigma w - \lambda w^T \mu \\
\text{subject to} \quad & w^T \mathbf{1} = 1
\end{aligned}
\tag{3.4}
$$

where λ is a parameter that allows to trade-off between the two objectives (this is the risk appetite in this case).

If we set $\lambda = 0$, we are minimizing the volatility of the portfolio and obtain the minimum variance portfolio. If we set $\lambda = \infty$, we are maximizing the return of the portfolio without taking into account the volatility of the portfolio.

The exact value of λ depends on the risk preference of the investor: the higher λ, the more higher the risk appetite.

Example X.1. *illustrates with various values of λ*

3.2 Climate Efficient Mimicking Portfolio

We can use the derived signal b to compute a constrained long-only portfolio ($w_i \geq 0$) and fully invested ($w^T\mathbf{1} = 1$) mimicking portfolio. We are now using the covariance matrix Σ of the assets to construct the *efficient mimicking portfolio*, along the signal b.

The optimization problem is:

$$\begin{aligned} \min_{w} \quad & \frac{1}{2}w^T\Sigma w - \lambda w^T b \\ \text{subject to} \quad & w^T\mathbf{1} = 1 \\ & w_i \geq 0 \end{aligned} \tag{3.5}$$

For the sake of simplicity, let's define $\Sigma = I$ (the identity matrix) and $\lambda = 2$. The minimization function becomes $w^T w - 2w^T b$. The solution of this problem is given by:

$$2w - 2b = 0 \Rightarrow w^* = b \tag{3.6}$$

This case corresponds to the intuitive weighting scheme we have seen in the previous chapter, with the weights defined as the signal b. Therefore, our previous case is a special case of the efficient mimicking portfolio when the covariance matrix is the identity matrix. In other words, **the signal b is an optimal weighting scheme only if the assets are uncorrelated** (*i.e.* the covariance matrix is the identity matrix).

Another interesting case is when $\lambda = 0$ (and $\Sigma = I$). In that case, the solution becomes $w^* = 1/N$. Therefore, **if the investor has absolutely no risk appetite, the equal-weighted portfolio is the optimal portfolio, regardless of the signal** b.

Example X.1. *illutrates with various values for* λ

3.3 Further Reading

Portfolio optimization: *Introduction to Risk Parity and Budgeting* by Roncalli (2013) [?]

Climate Efficient: *Factor-Mimicking Portfolios for Climate Risk* by De Nard *et al.* (2024) [?]

Chapter 4

Conclusion

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