Some Takeaways - Hedge and Tracking

Fernando Urbano

Autumn 2024

Attention: This is not the complete material. The following are organized notes from the lecture with key takeaways. We highly recommend that you also study the PowerPoint slides and supplement your understanding with your own lecture notes.

1 Net Exposure - Basis

An investor is long \$1 of "i" and hedges by selling \$h of "j".

The net exposure after hedging is known as the basis.

The position is perfectly hedged if:
$$P(\epsilon_t = 0) = 1$$

1.1 Why would we want to create a basis or hedge part of our position?

- **Non-tradable exposure**: If I own an asset that is not liquid and I want to hedge one particular risk that this asset carries, I could hedge, for example, the interest rate risk of a real estate investment while keeping some basis because the interest rate is not the only risk involved.

- **Liquid security but costly to short**: If shorting is expensive, we might hedge by shorting a related liquid security to avoid large fluctuations. For instance, I might own stocks in a small-cap tech company and short the S&P to hedge against market fluctuations.

2 Basis Risk

The basis is given by:
$$\epsilon_t = r_t^i - h r_t^j$$

We define the risk in the basis as the uncertainty related to the basis. More specifically, the basis risk generally refers to the volatility in ϵ_t, denoted by σ_ϵ:
$$\sigma_\epsilon^2 = \sigma_i^2 + h^2 \sigma_j^2 - 2h \sigma_i \sigma_j \rho_{i,j}$$

We can see that if $\rho_{i,j} = 1$ or $\rho_{i,j} = -1$, the basis risk is eliminated. Otherwise, we minimize the basis risk by adjusting h.

- **Higher correlation** implies a larger hedge ratio h, as the assets align more, making the hedge more effective.

- **Higher relative volatility of "i"** implies a larger h.

- **Negative correlation** requires going long on the hedging security.

The optimal hedge ratio that minimizes basis risk is:

$$h^* = \arg\min_{h} \sigma_\epsilon^2$$

Which simplifies to:

$$h^* = \frac{\sigma_i}{\sigma_j} \rho_{i,j}$$

2.1 Hedging Multiple Risks

When we have multiple regressors, the result becomes a regression where each asset representing a risk is a regressor.

$$r_t^i = \beta^{i,1} r_t^1 + \beta^{i,2} r_t^2 + \cdots + \beta^{i,k} r_t^k + \epsilon_t$$

The basis becomes the net return exposure, and the absolute value of the optimal hedge is determined by the betas in the return regression above (the hedge amount of each asset i is β^i).

The variances and correlations between assets change daily, so optimal hedges must adjust accordingly.

This applies to the excess returns of both the asset to be hedged and the hedging assets:

$$\tilde{r}_t^i = \beta^{i,1} \tilde{r}_t^1 + \beta^{i,2} \tilde{r}_t^2 + \cdots + \beta^{i,k} \tilde{r}_t^k + \epsilon_t$$

2.2 Should we include a constant (α) in the regression?

If I exclude an intercept, I get a truer picture of the situation. For example, when hedging tech stock with the S&P, I hedge with instruments. Including α in the regression adds a part of the hedge that cannot be bought — α cannot be "bought."

Why might it be reasonable to include an intercept? Including an intercept lets the betas focus on matching return variation, not just the level. If α is excluded, betas also adjust the magnitude. Including α ensures the hedge is more correlated with the returns.

The key point is whether we expect the difference in mean returns to persist out-of-sample. Historical data often provides uncertain estimates of mean returns, so excluding α may lead betas to "predict" differences in mean that aren't predictive of the future.

2.3 Should we do the regression in excess returns or returns?

Empirically, it doesn't matter much. Both approaches will yield similar results.

2.4 Investment Hedging a Factor

Hedging allows an investor to take a position on a specific thesis without exposing themselves to unwanted risks.

For example, if an investor believes in Uber's growth potential and buys its stock, they are exposed to market risk due to Uber's correlation with the market factor. By hedging with the S&P, the investor's performance depends only on Uber's relative performance versus r^m.

2.5 Properties of a Market-Hedged Position

When we hedge the market and invest in a particular asset, our return becomes $\alpha + \epsilon_t$:

$$\tilde{r}_t^i - \beta^{i,m} \tilde{r}_t^m = \alpha + \epsilon_t$$

In comparison to simply going long $\tilde{r}_t^i$, the strategy is no longer subject to the volatility from $\beta^{i,m} \tilde{r}_t^m$.

Investors expect α to be positive and ϵ_t to have low variance. If ϵ_t is sufficiently small, we may consider this a statistical arbitrage: a large α with low σ_ϵ.

You can also hedge multiple factors, with α and σ_ϵ maintaining similar properties.

3 Hedging vs. Tracking

In tracking, we don't invest in the left-hand side asset but aim to mimic its returns using the right-hand side assets. If we cannot invest in $\tilde{r}_t^i$, tracking allows us to buy assets on the right to mimic asset i's returns:

$$\tilde{r}_t^i = \alpha + \beta^{i,1} \tilde{r}_t^1 + \beta^{i,2} \tilde{r}_t^2 + \cdots + \beta^{i,k} \tilde{r}_t^k$$

In a hedging portfolio, we aim for:

- A large α (especially out-of-sample).
- A low σ_ϵ.

In a tracking portfolio (mimicking the left side), we aim for:

- $\alpha = 0$, to ensure good mimicry.
- A low σ_ϵ (for a high R^2).

For hedging, σ_ϵ is referred to as the basis. For tracking, it's called tracking error.

3.1   Practical Applications

- **Hedging**: Buy the left-hand side (asset i) and sell β dollars of the right-hand side factors.

- **Tracking**: Buy β dollars of the right-hand side factors to mimic the left-hand side (asset i), which we cannot buy.

The information ratio, α/σ_ϵ, measures the tradeoff between obtaining extra return α at the cost of ϵ.

For broad market factors, mutual funds often track factors, while hedge funds aim to hedge them. Both may use linear factor models to achieve this.

4   VaR (Value at Risk)

Value at Risk for quantile q is the return threshold where the probability of a return worse than this is q.

Common VaR quantiles are 1% and 5%.

4.1   Historical VaR

Historical VaR does not assume any distribution. It uses the empirical CDF (cumulative distribution function), meaning we derive the quantile directly from historical data.

4.2   Parametric VaR

Instead of sorting the data, we can assume a distribution and derive VaR from it.

Normal distribution is often assumed due to its simplicity in calculating VaR and CVaR.

For a normal distribution, VaR is:
$$r^{VaR_{q,\tau}} = \mu_\tau + z_q \sigma_\tau$$

Using the z-score, we can find the parametric VaR for any quantile.

4.3   Why Parametric VaR and CVaR May Be Preferred

Empirical estimation lacks statistical power compared to parametric estimation. When using empirical VaR, only a small sample portion is used to estimate the VaR for the desired quantile. Parametric approaches use the whole sample to calculate volatility, then estimate VaR.

This difference is more significant when we adjust the VaR period based on market changes. For example, if we focus on the last 20 days instead of five years, empirical VaR may rely on a single data point. In contrast, parametric methods estimate volatility over the 20 days, offering more statistically robust results.

Thus, while parametric methods may introduce bias (e.g., assuming normality), they provide better precision, particularly for small samples.

Calculating VaR with a normal distribution is also easier.

Generally, parametric VaR outperforms empirical VaR.

The same holds for CVaR: parametric CVaR provides more statistical significance than its empirical counterpart.

5   CVaR Formula

The analytical formula for CVaR is:

$$r_t^{CVaR_{q,\tau}} = \mu_{\tau,t} - \frac{\phi_z(z_q)}{q}\sigma_{\tau,t}$$

VaR and CVaR scale similarly to volatility. For instance, to annualize daily VaR, multiply by the square root of trading days in a year.

6   Hit Ratio

The hit ratio measures the percentage of times the next period's return (e.g., daily) is worse than the VaR estimate.

A hit ratio greater than q indicates that tail risk is worse than predicted by VaR of quantile q. A hit ratio smaller than q means VaR overestimates tail risk.

The ideal one-day VaR for the q percentile is a VaR with a hit ratio of q.

7   Hit Ratio Error

Hit Ratio Error is the hit ratio divided by the quantile used.

For example, if the hit ratio is 7% and the VaR corresponds to the 5th quantile, the hit ratio error is 40%.