Asset Allocation II

FINC-460 Investments

Kellogg School of Management

- How do we implement Mean-Variance Optimization Theory in practice?
- We need to select a set of assets to invest in. How?
- Once we selected the menu of assets, we need to obtain the inputs to the problem, i.e. means, variances and correlations.
- What do we do if some of the assumptions are violated?

Using the Markowitz model

- With a reasonable number of securities, the number of parameters that must be estimated is huge:
  - $\hookrightarrow$ For a portfolio of N (100) securities we need:

σ_i's	N	100
$E(r_i)$'s	N	100
$Cov(r_i, r_j)$'s	$\frac{1}{2}N(N-1)$	4950
Total	$\frac{1}{2}N(N+3)$	5150

  - $\hookrightarrow$ About how much data will we need when we have 500 securities? 1000 Securities?
- Means and covariances are estimated with *error*.
- Small errors in mean or covariance estimates often lead to unreasonable weights.

The top town approach

- In the US there are tens of thousands of traded securities. Clearly, we cannot solve Markowitz with these many assets.
- We can dramatically reduce the dimensionality of the problem by choosing across *portfolios* of assets.
- By specifying our choice set over a small number $(N \leq 50)$ of portfolios, we can find the optimal combination with the highest Sharpe ratio.
- Large institutional investors such as insurance companies and pension funds favor this approach, which they term 'top down'.

The top town approach

- How should we group securities into portfolios?
- We should group them based on an economically meaningful characteristic:
  - $\hookrightarrow$ Industry
  - $\hookrightarrow$ Geography
  - $\hookrightarrow$ Asset class

The top town approach

- Are there other advantages to portfolios rather than individual securities in Markowitz ?
- Portfolios already diversify some of the idiosyncratic risk
- This reduces estimation error in
  - $\hookrightarrow$ Expected returns
  - $\hookrightarrow$ Covariances
- Recall that measurement error in expected return of asset i equals $\sigma/\sqrt{T}$.
- Diversified assets have lower volatility σ, hence lower measurement error.

The top town approach

- Using portfolios rather than individual securities, allows us to use longer history of data.
  - ↪ Increasing T reduces measurement error.
- Individual firms may appear for a small part of the sample.
- However, as long as there exist similar firms, using portfolios we can use a much longer history
  - ↪ Example: we have data on portfolios of value/growth firms since 1926
  - ↪ How many firms today existed in 1926?

The top town approach

- Are there any disadvantages to not using individual securities?
- Our choice of which portfolios to focus on, might affect how well we will be able to do.
- Portfolios need to be different in economically meaningful ways.
- What will happen if you consider 26 portfolios, each grouping firms according to their first letter of their name?
- We can optimize over portfolios formed using any trading strategies.

Measurement error

- MV Optimization gives always the right answer, conditional on
  - ↪ the inputs
  - ↪ the assumptions
- Figuring out which inputs to use is going to be the hardest problem.
- Some of these inputs (e.g. average returns) are crucial, yet they are very difficult to estimate.
- Common practice is to use recent history to estimate these moments.
  - ↪ When deciding on how much data to use, we are trading off estimation error versus the likelihood that the past is not a good indicator of the future.

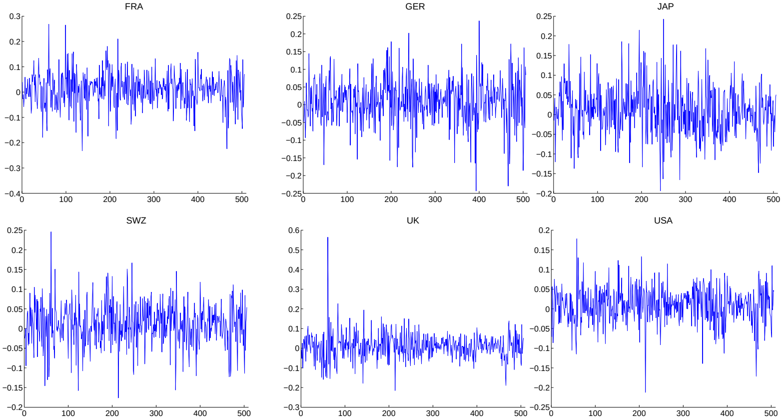
Example

- Consider the problem of investing in the following 6 international equity indices: **France, Germany, Japan, Switzerland, UK** and the **USA**.
- Download monthly historical index returns from MSCI.
- Compute historical mean and covariance matrix

$$\hat{\mu} = \begin{pmatrix} 1.02 \\ 1.00 \\ 0.94 \\ 1.06 \\ 1.03 \\ 0.88 \end{pmatrix} \quad \text{and} \quad \hat{\Sigma} = \begin{pmatrix} 44.0 & 30.9 & 17.9 & 23.8 & 26.5 & 17.0 \\ 30.9 & 41.8 & 16.3 & 24.6 & 22.1 & 16.0 \\ 17.9 & 16.3 & 39.0 & 15.2 & 16.4 & 10.0 \\ 23.8 & 24.6 & 15.2 & 28.8 & 21.3 & 13.7 \\ 26.5 & 22.1 & 16.4 & 21.3 & 41.8 & 17.0 \\ 17.0 & 16.0 & 10.0 & 13.7 & 17.0 & 20.5 \end{pmatrix}$$

- Feed the inputs into Markowitz, assume monthly risk-free rate of 5%/12

Return series



- We have ≈ 500 months of data, ranging from 1970 to 2012

Portfolio Optimization using Markowitz

Number of securities: 6

No	Name	Fraction	Expected Return	Standard Deviation
1	'FRA'	-2.03%	1.02%	6.6%
2	'GER'	-3.31%	1.00%	6.5%
3	'JAP'	16.64%	0.94%	6.2%
4	'SWZ'	49.78%	1.06%	5.4%
5	'UK'	7.16%	1.03%	6.5%
6	'USA'	31.76%	0.88%	4.5%
		100.00%		

Correlations		1	2
1		1.0	0.7
2		0.7	1.0
3		0.4	0.4
4		0.7	0.7
5		0.6	0.5
6		0.6	0.5

Corr OK? YES

Results:

Portfolio's Expected Return	0.0098
Portfolio's Standard Deviation	0.0434

Risk Free Rate 0.42%

Risk Aversion Coefficient: A= 10.00

Slope of CAL 0.1300

Weight on optimal risky portfolio: x*= 29.95%

Measurement error

- The MV-Optimizer returns the following set of weights

$$w^* = \begin{pmatrix} -0.020 \\ -0.033 \\ 0.166 \\ 0.498 \\ 0.071 \\ 0.318 \end{pmatrix}$$

- However, when we do MV-optimization, the computer thinks that the expected returns and covariance that we input is the truth (i.e. $\hat{\mu} = \mu$ and $\hat{\Sigma} = \Sigma$).
- Is there a way to quantify how far our estimated mean (or covariance) is from the truth?
- We can use statistics to build a confidence interval for the true mean return μ

Standard error of the mean

- A 95% confidence interval for the mean can be constructed as

$$\hat{\mu} \pm 1.96 \times s(\mu)$$

where $s(\mu) = \sigma/\sqrt{T}$, where T is the number of observations we used.

- In our case, the confidence interval is

$$CI = \begin{pmatrix} 1.02 \\ 1.00 \\ 0.94 \\ 1.06 \\ 1.03 \\ 0.88 \end{pmatrix} \pm 1.96 \times \begin{pmatrix} 0.295 \\ 0.287 \\ 0.278 \\ 0.238 \\ 0.287 \\ 0.201 \end{pmatrix}$$

- Not very comforting. And we did not even take the estimation error in the covariance matrix into account!

Sensitivity Analysis

- Monte Carlo simulations help quantify the effect of measurement error on our decision problem

  ↪ Estimate the mean and covariance of asset returns using available data

  ↪ Compute the MV-efficient portfolio

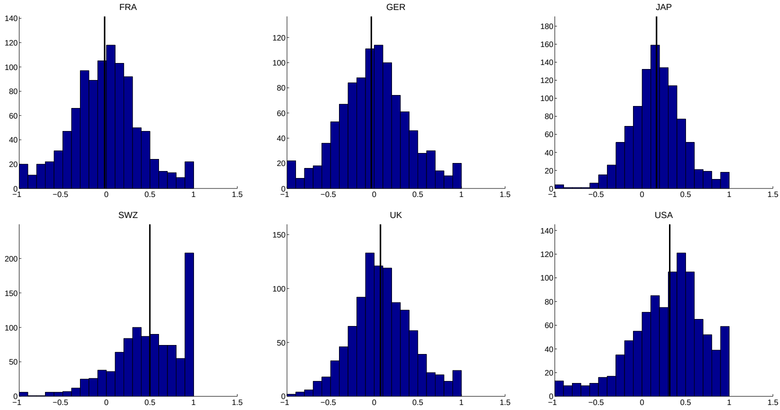
  ↪ Suppose that this is now the truth. Let's get into the brain of an investor who imperfectly observes this 'truth' and makes investment decisions.

  1. Simulate one realization of asset returns using the historical mean and covariance matrix.
  2. Estimate the mean and covariance matrix for this simulation. The investor in this simulation believes these are the right inputs to Markowitz.
  3. Feed these inputs into Markowitz – estimate weights of the optimal portfolio from the perspective of this investor.
  4. Simulate the performance of this 'optimal' portfolio using the 'true' moments of returns.

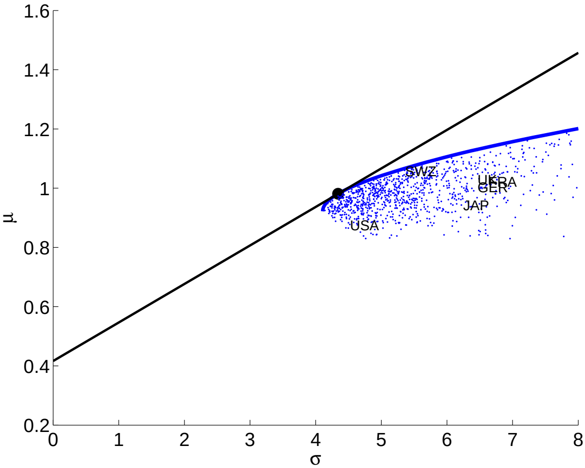
  ↪ Repeat (1)-(4) a large number of times

  ↪ Plot distribution of optimal portfolio weights across simulations

Sensitivity Analysis



- Weights restricted to $[-1, 1]$. Black line represents optimal allocation using historical moments. Histogram displays variability of portfolio weights across simulations.



- What we think is the MVE portfolio may be far from the truth.

Portfolio Optimization with Constraints

- There are two ways of getting around this problem

  1. use more conservative estimates for expected returns.
     - We could “shrink” the expected returns away the historical return towards some prior belief about the mean.
  2. impose constraints on the portfolio weights to ensure that we do not take extreme positions.

- We will revisit suggestion (1) when we see the Black-Litterman model. For now, lets examine (2).

Portfolio Optimization with Constraints

- Consider the previous problem, but now introduce portfolio constraints.
  - ↪ no short-selling
  - ↪ no more than 30% allocation in one country
- We can introduce these constraints into Markowitz.
- Constraints clearly shrink the set of feasible portfolios.
  - ↪ Hopefully it buys us some robustness instead.

Portfolio Optimization with Constraints

Number of securities: 6

No	Name	Fraction	Expected Return	Standard Deviation
1	'FRA'	1.44%	1.02%	6.6%
2	'GER'	5.00%	1.00%	6.5%
3	'JAP'	20.63%	0.94%	6.2%
4	'SWZ'	30.00%	1.06%	5.4%
5	'UK'	12.93%	1.03%	6.5%
6	'USA'	30.00%	0.88%	4.5%
		100.00%		

Correlations		1	2
1		1.0	0.7
2		0.7	1.0
3		0.4	0.4
4		0.7	0.7
5		0.6	0.5
6		0.6	0.5

Corr OK? YES

Results:

Portfolio's Expected Return	0.0097
Portfolio's Standard Deviation	0.0433

Risk Free Rate 0.42%

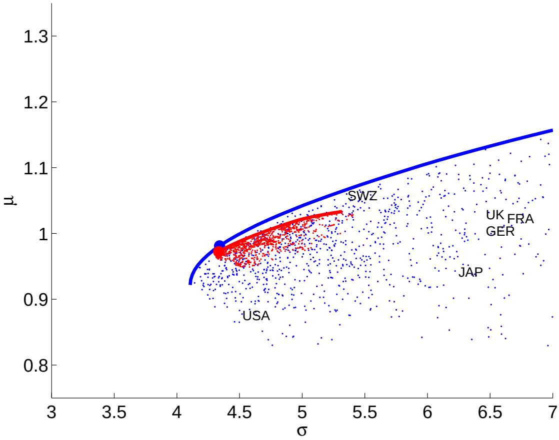
Risk Aversion Coefficient: A= 10.00

Slope of CAL 0.1282

Weight on optimal risky portfolio: x*= 29.59%

- Imposing short sales lowers our Sharpe ratio.
- But may lead to more robust estimates

Portfolio Optimization with Constraints



- Constrained frontier (red) lies inside the unconstrained frontier (blue)
- But the performance of the 'in-sample' MVE portfolio across simulations is closer to the frontier

- Portfolio constraints are one way to prevent Markowitz from 'reading too much into historical data'.
- More generally, there may be other reasons why you would want to limit exposure to certain classes of assets.
    - i) Transaction costs.
    - ii) Liquidity.
    - iii) Non-tradeable income
- We should recognize that often some of the assumptions behind MV-Optimization are violated. Imposing portfolio constraints can limit potential problems.

Portfolio Optimization with Non-Tradeable Assets

- One of the assumptions behind MV Optimization is that we are optimizing our allocation over **all** sources of wealth.
- This assumption is clearly unrealistic. In practice there are many sources of income that we cannot convert into a tradeable asset
  - ↪ Human capital (present value of human capital)
  - ↪ Income from operations
  - ↪ Alumni donations
  - ↪ Pension contributions
- Other times we own assets that we could trade, but that act as more than investment (e.g. housing).
- We can embed these types of constraints into Markowitz by introducing new assets that we cannot trade.

Portfolio Optimization with Non-Tradeable Assets

- Let's return to the previous problem, but now suppose that we are working for a financial institution that has extended a large number of loans to Swiss banks.
- The value of this loan portfolio accounts for 30% of our total assets.
- For relational reasons, we do not want to sell this loan portfolio.
- How should this affect our equity allocation decision?
- We need to make some assumptions: This loan portfolio
  - $\hookrightarrow$ has an expected return of 0.3% monthly and a standard deviation of 2%.
  - $\hookrightarrow$ has a correlation of 70% with the Swiss equity index and 40% with everything else.

Portfolio Optimization with Non-Tradeable Assets

Number of securities: 7

No	Name	Fraction	Expected Return	Standard Deviation
1	'FRA'	0.13%	1.02%	6.6%
2	'GER'	-1.35%	1.00%	6.5%
3	'JAP'	10.61%	0.94%	6.2%
4	'SWZ'	35.84%	1.06%	5.4%
5	'UK'	8.14%	1.03%	6.5%
6	'USA'	16.63%	0.88%	4.5%
7	LOANS	30.00%	0.30%	2.0%
		100.00%		

Correlations		1	2
1		1.0	0.7
2		0.7	1.0
3		0.4	0.4
4		0.7	0.7
5		0.6	0.5
6		0.6	0.5
7			

Corr OK? YES

Results:

Portfolio's Expected Return	0.0079
Portfolio's Standard Deviation	0.0355

Risk Free Rate	0.42%	**Risk Aversion Coefficient: A=**	10.00
Slope of CAL	0.1041	**Weight on optimal risky portfolio: x*=**	29.30%

- How does our portfolio change relative to the previous case. Why?

Portfolio Optimization with Non-Tradeable Assets

- Where can we get the inputs for non-tradeable assets?
- Sometimes there are comparable assets that are traded in the market (e.g. loans vs bonds).
- But often it is difficult to do without a particular model (e.g. human capital).

Investment Horizon

- Mean-Variance analysis is a static problem.
  - $\hookrightarrow$ It can be applied to any horizon.
- The choice of horizon merely affects the inputs.
  - $\hookrightarrow$ i.e. if have a monthly horizon, using monthly returns to estimate inputs
- In practice, choice of horizon will not affect the expected returns
- But, it may play a role when estimating risk
  - $\hookrightarrow$ finance textbooks mention the square-root of time rule:

$$\sigma(\text{K-period return}) = \sqrt{K} \times \sigma(\text{1-period return})$$

  - $\hookrightarrow$ What is the implicit assumption here?

Investment Horizon

- Our portfolio allocation formula says stock allocation should depend on the ratio of mean to variance of total return:

$$w = \frac{1}{A}\frac{E(R) - r_f}{\sigma^2}$$

- Should an investor with a 10 year horizon allocate more to stocks than an investor with a 1 year horizon because "stocks are safer in the long run" and he can "wait out market declines"?

- What does the answer depend on?

- Why would it be a mistake to look at annualized returns to answer this question?

Risk and Investment Horizon

Horizon	$var(R_H)$	$\frac{var(R_H)}{var(R_1)}$
1-month	0.003	
1-year	0.040	13.54
2-year	0.086	28.82
3-year	0.117	39.34
4-year	0.147	49.17
5-year	0.165	55.47
10-year	0.240	80.63

- Table shows variance of H-period returns on market portfolio since 1927
- What is going on here? Does the finance textbook give the right advice?
  - $\hookrightarrow$ Which assumption that we made is wrong?
- In practice, also the matrix of correlations changes with the horizon...

Estimating Risk

- Another issue that becomes important when you manage large portfolios is obtaining accurate estimates of the covariance matrix.
- The number of parameters you need to estimate for N assets is $N(N-1)/2$. It grows very fast as N increases.
- In same cases, we can simplify the problem dramatically by assuming that a small number of common factors drives the correlation of asset returns.

A single index model

- An alternative to estimating $\sigma_{i,j}$'s is to assume that a *single index model* describes returns:

  1. returns can be decomposed into a 'systematic' and a 'non-systematic' part

  $$\tilde{r}_{i,t} - r_{f,t} = \alpha_i + \beta_i(\tilde{r}_{m,t} - r_f) + \tilde{\varepsilon}_{i,t}$$

  2. the non-systematic part is truly asset-specific: for two different securities i and j,

  $$cov(\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{j,t}) = 0.$$

- When is this a reasonable simplification?
- We will later generalize this to K common factors.

A single index model

- To get covariances/correlations, use:

$$\sigma_{i,j} = cov(r_{i,t}, r_{j,t}) = \beta_i \beta_j \sigma_m^2 \quad \forall i \neq j$$

- For variances:

$$\sigma_{i,i} = \sigma_i^2 = \beta_i^2 \cdot \sigma_m^2 + \sigma_{\varepsilon,i}^2$$

$\hookrightarrow$ Notice that variances are unaffected!

- Then, the correlation between two asset returns i and j is

$$\rho_{i,j} = \frac{\beta_i \beta_j \sigma_m^2}{\sigma_i \sigma_j}$$

A single index model

- Notice now that the number of parameters is much fewer.
- Instead of $N \times (N-1)/2$ we only need
  - $\hookrightarrow$ K variances for each factor
  - $\hookrightarrow$ $N \times K$ betas
- For small number of factors (i.e. here $K = 1$) the reduction in parameters is quite dramatic

A single index model

- Where do the betas come from?
- They can be estimated via regression. We will revisit this issue again.
- For now, just recall the formula for the beta of asset i with portfolio p

$$\beta_{i,p} = \frac{cov(R_i, R_p)}{var(R_p)}$$

A single index model, example

- Let's go back to our international equity example. Suppose that there is one common factor, the 'global' market index, here defined as an equally weighted average of the six country portfolios.

- Then

$$var(R_G) = 22.1$$

and

$$\beta = \begin{pmatrix} 1.211 \\ 1.147 \\ 0.869 \\ 0.963 \\ 1.098 \\ 0.712 \end{pmatrix} \quad \text{imply that} \quad C_1 \begin{pmatrix} 44.0 & 30.6 & 23.2 & 25.7 & 29.3 & 19.0 \\ . & 41.8 & 22.0 & 24.3 & 27.7 & 18.0 \\ . & . & 39.0 & 18.4 & 21.0 & 13.6 \\ . & . & . & 28.8 & 23.3 & 15.1 \\ . & . & . & . & 41.8 & 17.2 \\ . & . & . & . & . & 20.5 \end{pmatrix}$$

is the covariance matrix, assuming the single index model is true.

A single index model, example

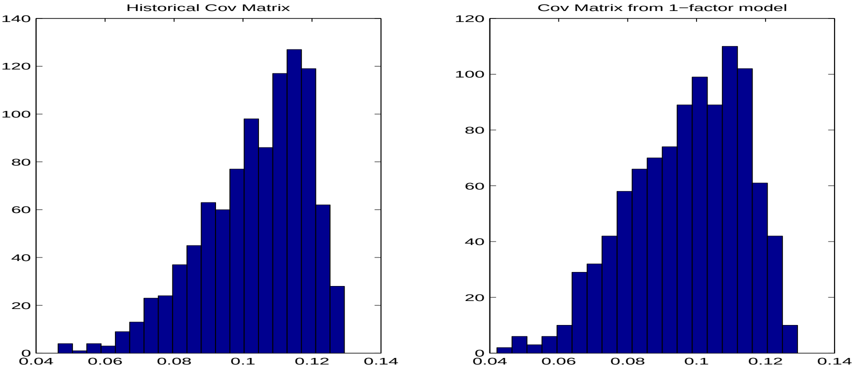
- Let's compare this with the case where we just using the historical covariance matrix $\hat{\Sigma}$

- Scaling into correlations,

$$\hat{C} = \begin{pmatrix} 1 & 0.72 & 0.43 & 0.67 & 0.62 & 0.57 \\ . & 1 & 0.40 & 0.71 & 0.53 & 0.55 \\ . & . & 1 & 0.45 & 0.41 & 0.35 \\ . & . & . & 1 & 0.61 & 0.56 \\ . & . & . & . & 1 & 0.58 \\ . & . & . & . & . & 1 \end{pmatrix} \quad \text{versus}$$

$$C_1 = \begin{pmatrix} 1 & 0.71 & 0.56 & 0.72 & 0.68 & 0.63 \\ . & 1 & 0.54 & 0.70 & 0.66 & 0.62 \\ . & . & 1 & 0.55 & 0.52 & 0.48 \\ . & . & . & 1 & 0.67 & 0.62 \\ . & . & . & . & 1 & 0.59 \\ . & . & . & . & . & 1 \end{pmatrix}$$

A single index model, example



- Graphs show distribution of realized Sharpe-ratios across simulations.
- Not much difference here. Reason is that with a large number of observations, covariance matrix is precisely estimated.
- Results are different when the number of assets is large...

A single index model

- The single index model can also help us make sense of a non-stationary environment.

- A common belief among market practitioners is that in times of high uncertainty (market volatility), all the asset correlations tend to increase.

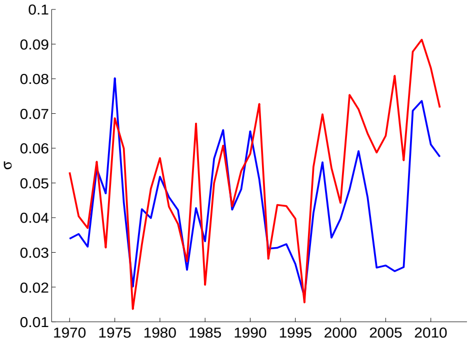
- Can we make sense of this statement using the 1-factor model?

  $\hookrightarrow$ Recall that

$$\rho_{i,j} = \frac{\beta_i \beta_j \sigma_m^2}{\sigma_i \sigma_j}$$

  $\hookrightarrow$ all else equal, an increase in the volatility of the common factor σ_m^2 should lead to an increase in the pairwise correlations of assets positively related to the common factor ($\beta > 0$)

Average correlations versus volatility of common factor



- Blue line plots the volatility of the common factor; red line plots the average correlation across the 6 equity indices.
- In both cases, I use a one-year window to estimate the volatility and correlation.

Uncertainty and correlations

- A common reason why MV theory fails in practice is that the world is non-stationary.
  - ↪ in other words, the past is not always a good predictor of the future.
- Furthermore, as correlations increase, the benefits of diversification disappear dramatically.
  - ↪ Estimating correlations from recent history can lead to mistaken beliefs about the benefits of diversification.
  - ↪ Case in point: $CDOs$ and $CDOs^2$.

Portfolio variance and correlations

- Consider a portfolio of $N = 100$ assets, each with $\sigma = 0.1$.
- Suppose that we form a equal-weighted portfolio of these assets.
- The variance of this portfolio is equal to

$$\sigma_p^2 = \left(\frac{1}{N^2}\right)\sum_{i=1}^{N}\sigma_i^2 + \sum_{i=1}^{N}\left(\frac{1}{N^2}\right)\sum_{\substack{j=1 \\ i\neq j}}^{N}\sigma_i\sigma_j\rho_{ij}$$

$$\left(\frac{1}{N}\right)\sigma^2 + \sigma^2 \times \bar{\rho}$$

i.e. it mostly depends on the average correlation across these securities

- Underestimates of correlation lead to underestimates of portfolio risk

A single index model

- The single index model allows us to make forecasts of the entire correlation matrix,
  1. assuming we have accurate measures of the risk exposures β
     - We will revisit this issue in Lecture 4
  2. and a forecast of the volatility of the common factor σ_M
- There a number of statistical techniques used to forecast variance (GARCH etc)
- We can also use our own views, or the implied volatility from options prices (see Derivatives class)
- **Bottom line:** A single factor model allows us to forecast the entire correlation matrix using a small(er) number of inputs

Conclusion

- The quality of the output in MV Analysis depends crucially on the quality of the inputs.
- Expected returns are difficult to estimate using historical data, yet they are very important in the analysis.
- One way to deal with the problem is to impose constraints on portfolio weights. Such constraints may be adhoc, but they may lead to more robust outcomes.
- In the coming lectures we will a way to form some a priori views on what the expected returns *should* be in equilibrium.
- Deviations from this equilibrium level of average return will lead to 'good deals'.