# Default Correlations

**derived with an averaging model**

FORTIS

FORTIS

# Abstract

In the estimation of the credit loss distribution of a portfolio, the correlation of default between obligors has a significant impact, increasing the bank Economic Capital requirements. Since it is practically difficult to measure default correlations, they are commonly inferred from default probabilities of the obligors and correlation of the underlying assets of the obligors. Moreover, asset return data is of higher quality and availability than credit default data. In this talk we discuss an averaging model that allows to create inter and intra default correlation between groups of similar clients from a large representative data set. This approach assumes that all the elements in the correlation matrix for a portfolio can be approximated by the average correlation of their peer groups in the data set. The similarity is based on a hierarchical clustering according to region, sector, rating and asset size. We describe the method from a practical perspective and discuss results.

Author: Luc.Hoegaerts@gmail.com,

Senior Risk Modeler, Fortis Central Risk Management, Belgium

FORTIS

# Table of Contents

# 1. Introduction
## 1.1. Basic elements of credit risk

Credit Risk is the risk that a borrower will be unable to pay back his loan.

For any individual contract, the future loss (in a one year period) is random, i.e. unknown in advance. In default models, such loss is commonly broken down in 3 factors:

1. Exposure-At-Default (EAD): amount at risk at the point of default, maximal (monetary) amount of loss on risk i, given that default occurs.

2. Loss-Given-Default (LGD) : degree of security of risk i, expected percentage of the EAD that will be lost, given that default occurs

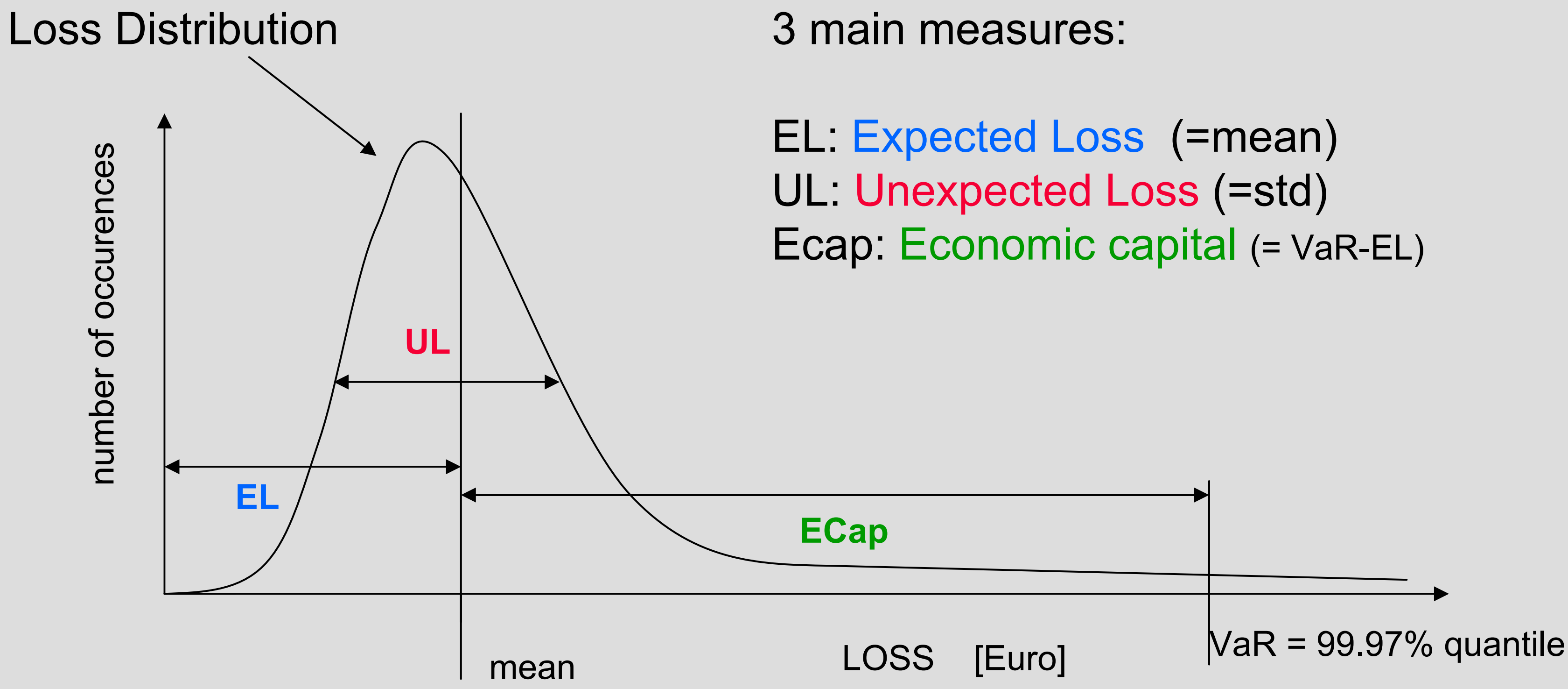
3. Probability of Default (PD): likeliness of default over e.g. 1 year.

Imagine portfolios where n clients default according to their specific parameters, then we get a series of each time an *aggregated loss S* per portfolio.

By counting occurences we can determine the probability per possible aggregated loss amount, which is the portfolio "loss distribution".

# 1.2. Quantifying Credit Risk

A bank can quantify its portfolio credit risk through the measurement of the expectation and variability of the portfolio credit loss. Fortis holds capital to protect against the volatility, such that it covers the minimal potential loss amount suffered in the 3 worst loss cases when considering 10,000 portfolio loss events (each time over 1 year).




3 main measures:

EL: Expected Loss (=mean)
UL: Unexpected Loss (=std)
Ecap: Economic capital (= VaR-EL)

# 1.3. Linear default correlations

For the portfolio EL, it suffices to add the stand-alone expected losses because

$$EL(X+Y+Z)= L(X) + E(Y) + E(Z)$$

However, for the portfolio UL, adding the stand-alone variances is not enough information because

$$var(X+Y+Z) = var(X) + var(Y) + var(Z) + 2\ \rho_{XY}\ + 2\ \rho_{XZ}\ + 2\ \rho_{YZ}$$

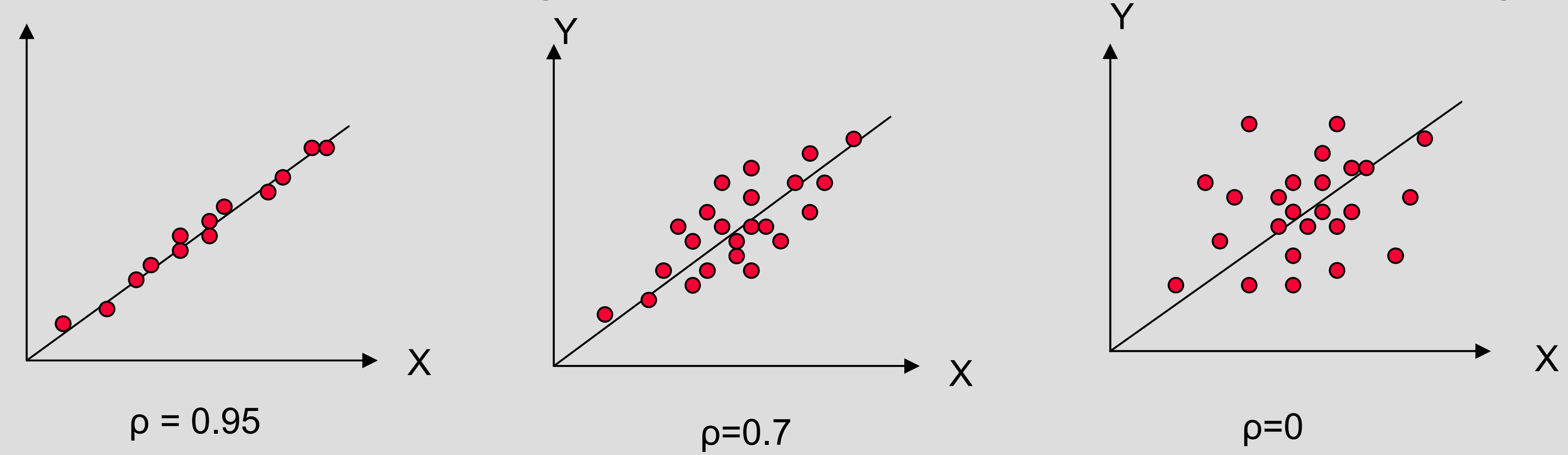
Hence: the linear default correlation $\rho$ between pairs is needed!

Research on real default data has shown that $\rho$ for sure is not negligable.

Default correlations become therefore an additional basic element of credit risk.

# 1.4. Correlation

Correlation indicates the degree in which observations deviate from a straight line




Correlation ρ is a measure of ***linear*** dependency between two obligors,
it indicates 'how much the time of default of Y is coincident with the one of X'

- If two obligors X and Y are statistical independent => $\rho = 0$.
- Range: $-1 \le \rho \le 1$ (sign indicates direction, value indicates magnitude)

FORTIS

# 1.5. Obtaining default correlations

Problem: defaulting pairs are unlikely to be observed, few data points...

$\Rightarrow$ We need to measure default correlations indirectly

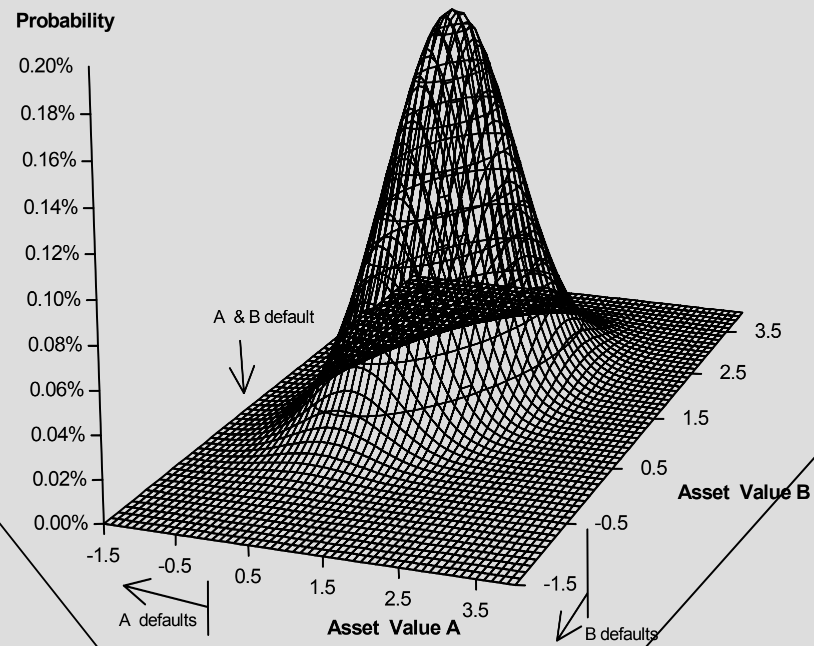
Solution: Asset correlations approximate strength of co-movement in time


$\rho_{Asset}$ is a building block for determining joint default probability

FORTIS

# 1.6. Joint default probability

Probability of Default $PD_A$ $PD_B$




asset correlation $\rho_{Asset}$

**Formula 1**

**proportion of volume under part of asset value surface where both hit default**

**Assumptions**:

- $PD_A$, $PD_B$ ***Bernoulli distributed***
- $JPD_{AB}$ ***standard normal distributed***

**$JPD_{AB}$**

Merton framework:

default when assets value falls below liabilities value

FORTIS

# 1.7. Default probability approximated

Logically, the default correlation is on its turn determined by the likeliness of default of the obligors apart (PD), but also on the likeliness that they default together (JPD):

**Formula 2:**

$$\rho_{AB} = \frac{\mathbf{JPD_{AB}} - PD_A\, PD_B}{\sqrt{(PD_A\,(1-PD_A) PD_B\,(1-PD_B))}}$$

**Conclusion:**

$$\rho_{Asset} \overset{1}{\Rightarrow} JPD \overset{2}{\Rightarrow} \rho_{default}$$

FORTIS

# 2. The averaging model

## 2.1. A basic approach

**Strategy**: asset correlations -> default correlations

*Ideally*: asset correlations needed between ***every*** pair of two clients

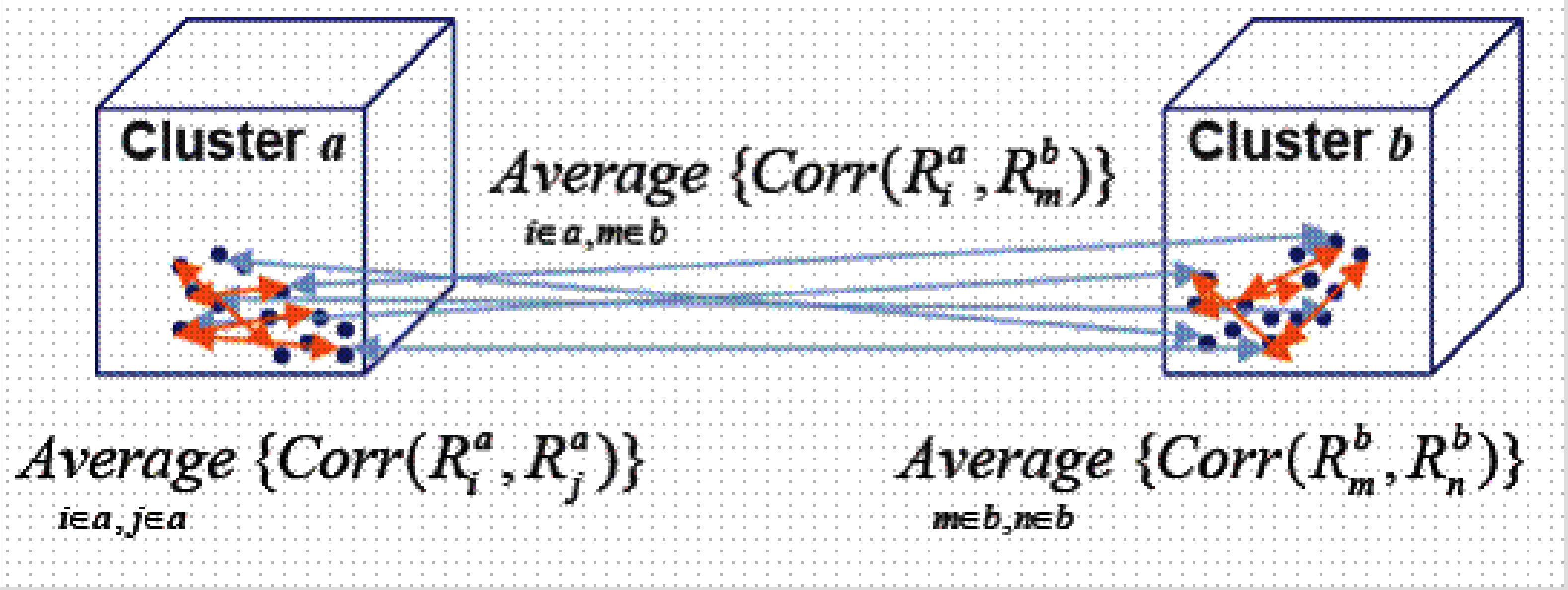
**Practical problem**: often no data of assets available of each client

**Solution**: Group similar clients together into **clusters** and assume that

- Assets of all clients within the same cluster are similarly correlated
- Correlations between asset co-movements from one cluster and another are equal

# 2.1. Taking averages over all combinations

- The intra-cluster correlation between any two firms in the same cluster is calculated by taking the average of all pair-wise correlations within the cluster.
- The inter-cluster correlation between any two firms in different clusters is calculated by taking the average of all pair-wise correlations between the two clusters.




**Next question**: How to group clusters? Depends on data availability!

FORTIS

# 2.3. Creation of the clusters

KMV Credit Monitor

- 64 843 companies
- over109 months (from 31/3/1997 until 30/4/2006)
- country (71), sectors (61), EDF, asset value

Natural asset cluster formation and mapping yields a data trade-off:

1. Geographical regions (3) [EUR,US,ASIA]
2. Sectors (7)
3. Client rating colors (4)
4. Asset size (4)

# 2.4. Sector division

- Group 1: Construction/Manufacturing (Hard)
- Group 2: Manufacturing (Soft)
- Group 3: Transport Manufacture, Communications, Utilities
- Group 4: Wholesale Trade and Distribution
- Group 5: Retail Trade and Sales
- Group 6: Financial Services
- Group 7: Services and Human Resources

FORTIS

# 2.5. Client rating colors

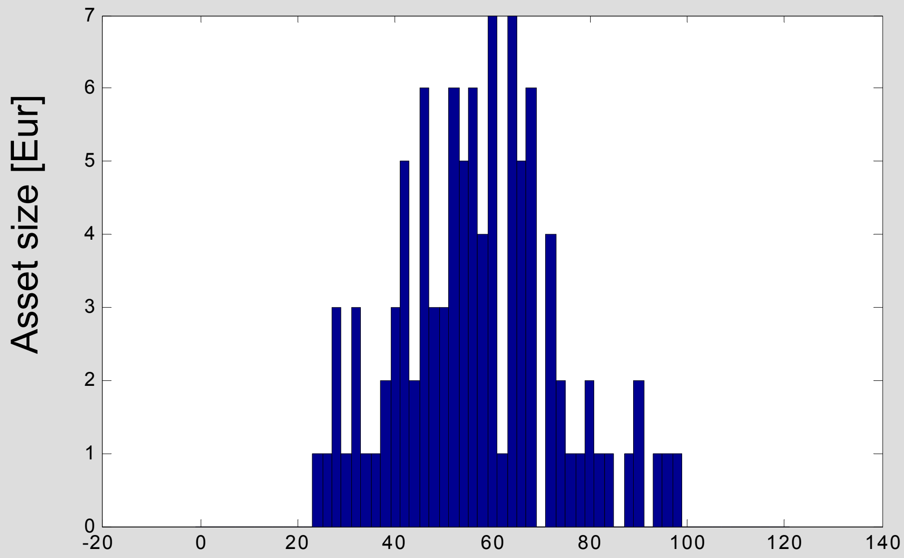
## Fortis New Masterscale

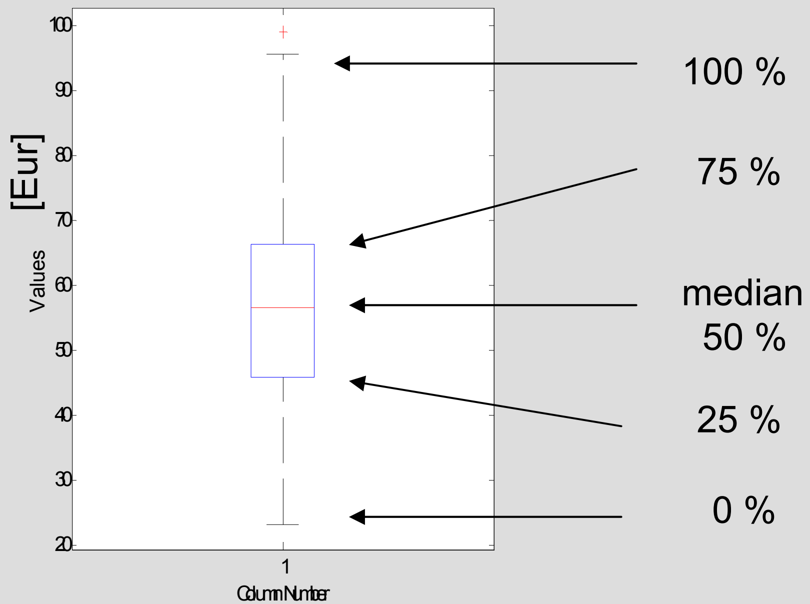
Mapping applicable as per 01/01/2006

| DEFAULT | Basel Rating | Probability of Default (PD) - % | | | S&P Ratings | Moody's Ratings | Banks | Sovereigns |
|---|---|---|---|---|---|---|---|---|
| | **0** | 0 | | | | | | |
| | **1** | 0,00 **(0,03)** 0,06 | 1.1 | 0,00 **(0,01)** 0,02 | AAA/AA+/AA/AA- | Aaa/ Aa1/Aa2/Aa3 | 1/2.1/2.2 | 2.1 |
| | | | 1.2 | 0,02 **(0,03)** 0,04 | A+ | A1/A2 | 2.3/3.1 | 2.2 |
| | | | 1.3 | 0,04 **(0,05)** 0,06 | A | A3 | 3.2 | 2.3/3.1 |
| | **2** | 0,06 **(0,09)** 0,12 | 2.1 | 0,06 **(0,07)** 0,08 | A- | | | |
| | | | 2.2 | 0,08 **(0,09)** 0,10 | | | 3.3 | 3.2 |
| | | | 2.3 | 0,10 **(0,11)** 0,12 | | Baa1 | | |
| | **3** | 0,12 **(0,16)** 0,20 | 3.1 | 0,12 **(0,13)** 0,14 | BBB+ | | 4.1 | 3.3 |
| | | | 3.2 | 0,14 **(0,16)** 0,17 | | | | |
| | | | 3.3 | 0,17 **(0,19)** 0,20 | | Baa2 | 4.2 | 4.1 |
| | **4** | 0,20 **(0,25)** 0,30 | 4.1 | 0,20 **(0,22)** 0,23 | | | | |
| | | | 4.2 | 0,23 **(0,25)** 0,26 | BBB | | | |
| | | | 4.3 | 0,26 **(0,28)** 0,30 | | | 4.3 | 4.2 |
| | **5** | 0,30 **(0,37)** 0,44 | | | | Baa3 | 5.1 | 4.3 |
| | **6** | 0,44 **(0,53)** 0,62 | | | BBB- | | | 5.1 |
| | **7** | 0,62 **(0,74)** 0,85 | | | BB+ | Ba1 | 5.2 | 5.2 |
| | **8** | 0,85 **(1,01)** 1,16 | | | BB | | 5.3 | 5.3 |
| | **9** | 1,16 **(1,37)** 1,58 | | | | Ba2 | 6.1 | |
| | **10** | 1,58 **(1,87)** 2,15 | | | BB- | | | 6.1 |
| | **11** | 2,15 **(2,54)** 2,92 | | | | Ba3 | 6.2 | 6.2 |
| | **12** | 2,92 **(3,45)** 3,97 | | | B+ | B1 | 6.3 | 6.3 |
| | **13** | 3,97 **(4,69)** 5,40 | | | B | | | 7.1 |
| | **14** | 5,40 **(6,38)** 7,35 | | | | B2 | | 7.2 |
| | **15** | 7,35 **(8,68)** 10,00 | | | B- | B3 | 7.1 | |
| | **16** | 10,00 **(12,00)** 14,00 | | | | | 7.2 | 7.3 |
| | **17** | 14,00 **(17,00)** 20,00 | | | CCC/C | Caa/C | 7.3 | |
| | **18** | Default - without LLR | | | Default | Default | Default | Default |
| | **19** | Default - with LLR | | | Default | Default | Default | Default |
| | **20** | Default - doubtful | | | Default | Default | Default | Default |

19/12/2005

# 2.6. Asset division




After looking for (i) region, (ii) sector, (iii) rating we arrive at a set of clients that we further need to subdivide into final clusters according to (iv) asset size:

Rule: Take the 0 – 25 – 50 – 75 – 100 % clients and set the boundaries correspondingly

# 2.7. Corporates

Selected firms out of KMV data except of public firms and individuals

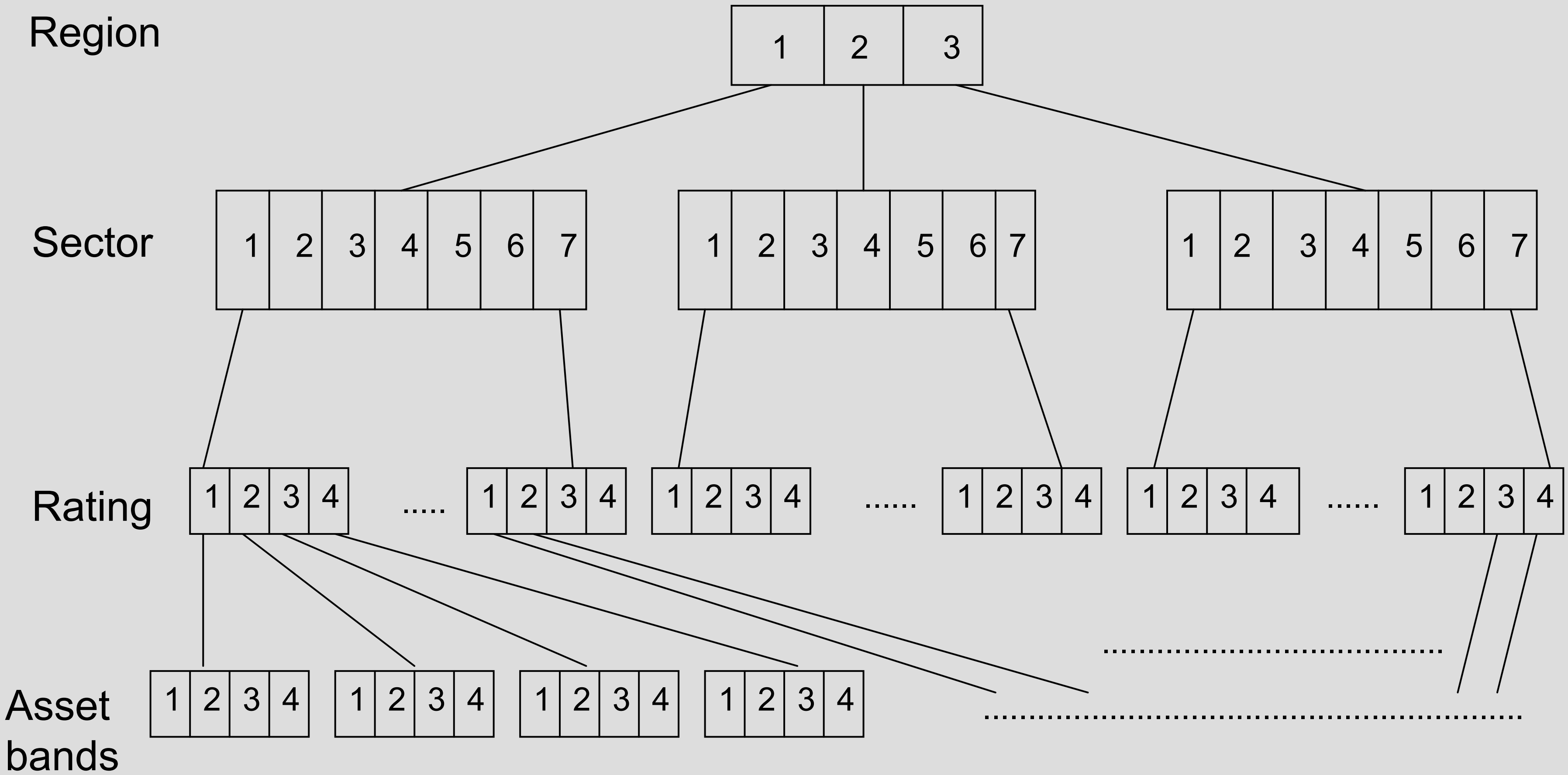
1. Region: firms from all 3 world regions
2. Sectors: firms from all 7 sectors
3. Rating: firms from all 4 client rating colors
4. Assets: firms from all 4 asset size bands

**In total**: 336 asset clusters

FORTIS

# Visualization

Corporates




In total we hence obtain 3x7x4x4 = 336 asset clusters

FORTIS

# 2.8. Public firms

Public firms are selected from KMV (except of privates) if

1. Region: all three regions
2. Sectors: belonging to the 5 KMV sectors
   - N02 Agriculture
   - N19 Education and Repair
   - N58 Water Supply
   - N59 Electricity Distribution
   - N60 Gas Management
3. Rating: Masterscale = 1 or 2
4. Assets > 7.5 billion Euro

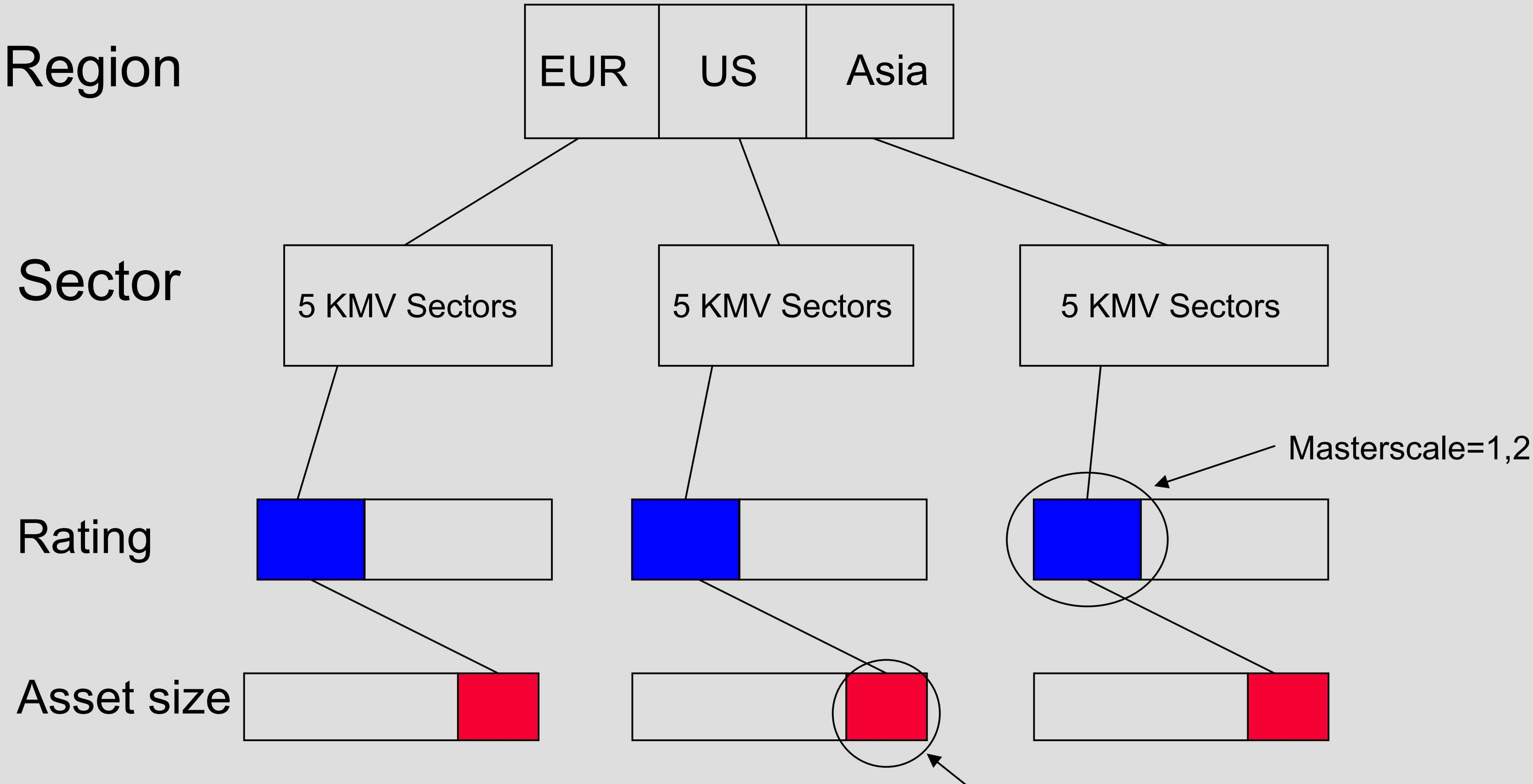
**In total**:

3 asset clusters (separate asset correlation cluster for each region)

FORTIS

# Visualization

Public firms




**In total**: **3 asset clusters**

FORTIS

# 2.9. Individuals

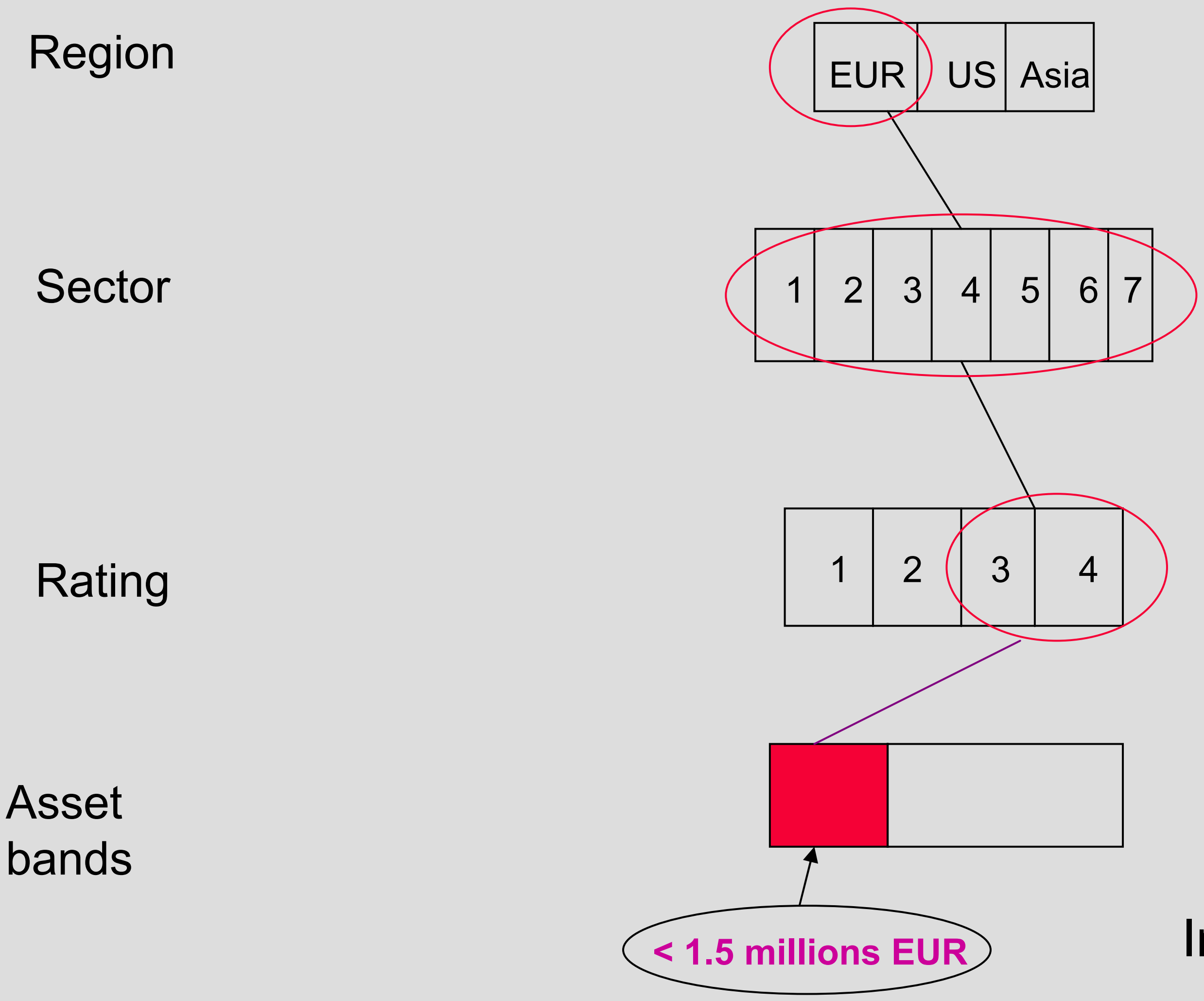
Selected out of KMV data if

1. Region: only Europe
2. Sectors: belonging to 14 particular KMV sectors
3. Rating: Masterscale rating ≥ 9
4. Assets: having less than 1.5 million euros in assets

**In total**: 1 asset cluster

# Visualization

Individuals




In total: **1 asset cluster**

FORTIS

# 2.10. Remarks on the cluster definition

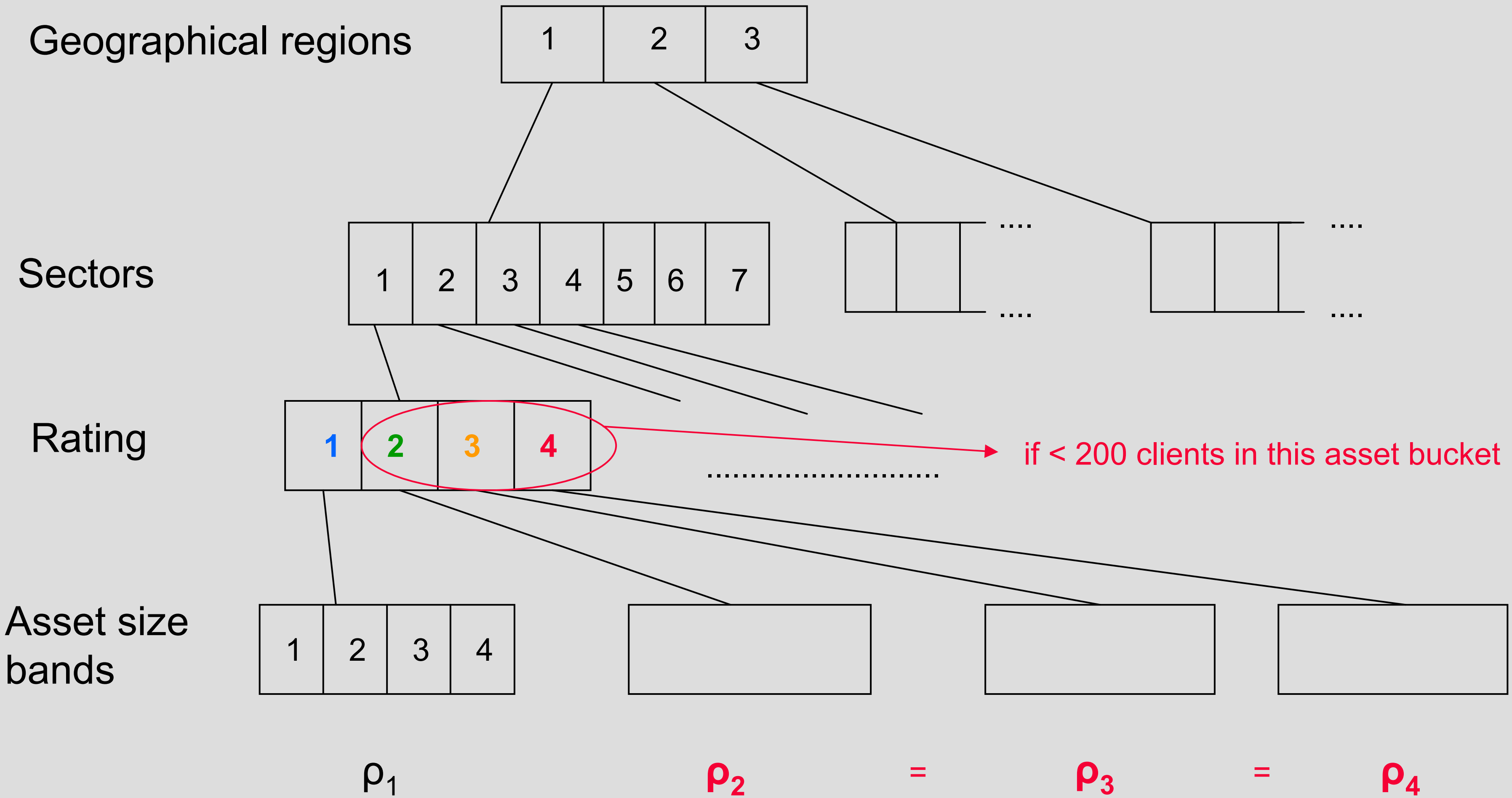
- Clusters with few data would create more uncertainty with respect to the stability of the asset correlations;

- Proposed solution: Group clusters with limited data together

- Practice: Grouping the corporate data on region, sector group and color if there were less than 200 companies at this level of grouping, then the 2nd, 3rd and 4th client rating colors were combined together.

- Result: 218 effective asset clusters.

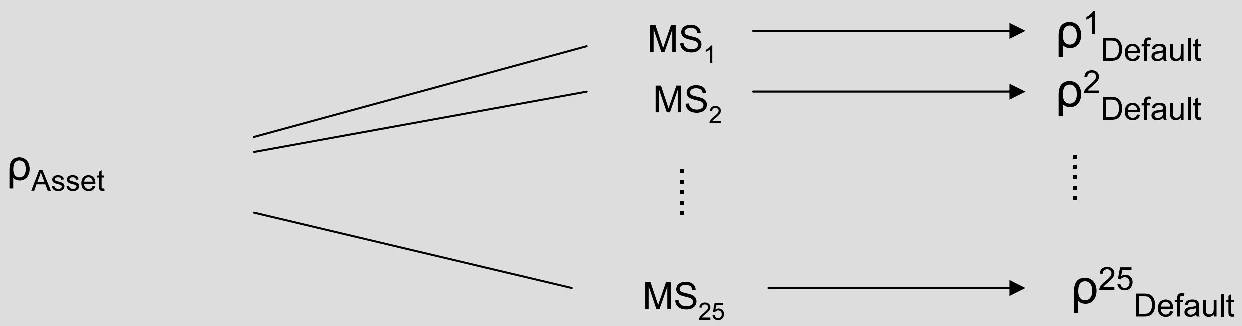
# Visualization:
# Grouping less populated clusters together



FORTIS

# 3. Default Correlation clusters

Principle: Make use of all 25 Masterscale values




•Corporates: same region, sector and asset size bands are kept but combined with all 25 Masterscale values

=> 3 x 7 x 4 x 25 = 2100 Default clusters

•Publics: The 3 regions are combined with the 25 Masterscale values

=> 3 x 1 x 1 x 25 = 75 Default clusters

•Individuals: Use all 25 Masterscale values =>1 x 1 x 1 x 25= 25 Default clusters

**Result**: 2100 + 75 + 25 = 2200 Default clusters

FORTIS

# 4. The single factor model

(1) Four-dimensional clusters are defined as previously.

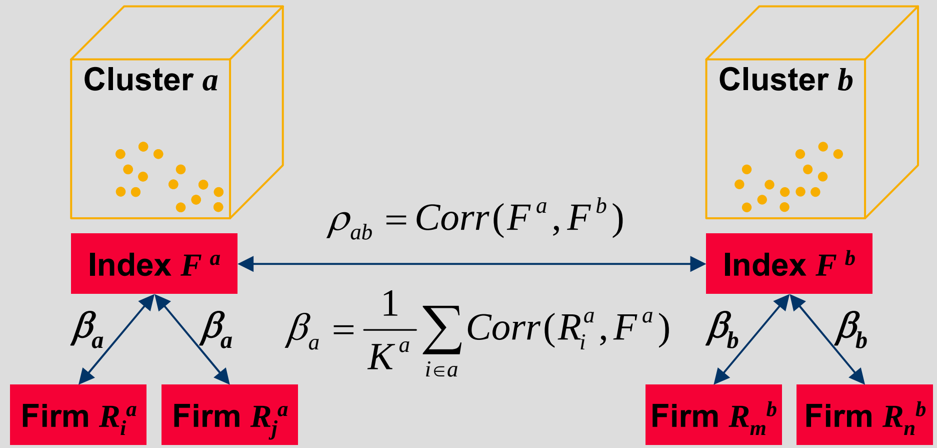
(2) Each firm is uniquely assigned to one cluster and an index $F^a$ is constructed (based on unweighted log asset returns).

(3) The index-to-index correlation $\rho_{ab}$ between any two indices $a$ and $b$ is calculated. For 340 clusters, this amounts to ≈58,000 index correlations.

(4) The firm-to-index correlation $\beta_a$ between any firm and its index is calculated by averaging over individual firm-to-index correlations. For ≈ 65,000 firms in total, this amounts to 65,000 correlations.

(5) The intra-cluster correlation between any two firms in the same cluster $a$ is calculated as $\beta_a^2$.

(6) The inter-cluster correlation between any two firms in different clusters $a$ and $b$ is calculated as $\beta_a \beta_b \rho_{ab}$ .




Each firm's standardized log asset return is assumed to follow a factor model of the form

$$R_i^a = \beta_a F^a + \sqrt{1-(\beta_a)^2} \cdot \varepsilon_i^a \quad with \quad F^a \sim N(0,1),\ \varepsilon_i^a \sim N(0,1)$$

Idiosyncratic shocks $\varepsilon_i$ are assumed independent of the systematic factor $F^a$.

**Intra-cluster** and **inter-cluster** correlations can then be estimated as follows

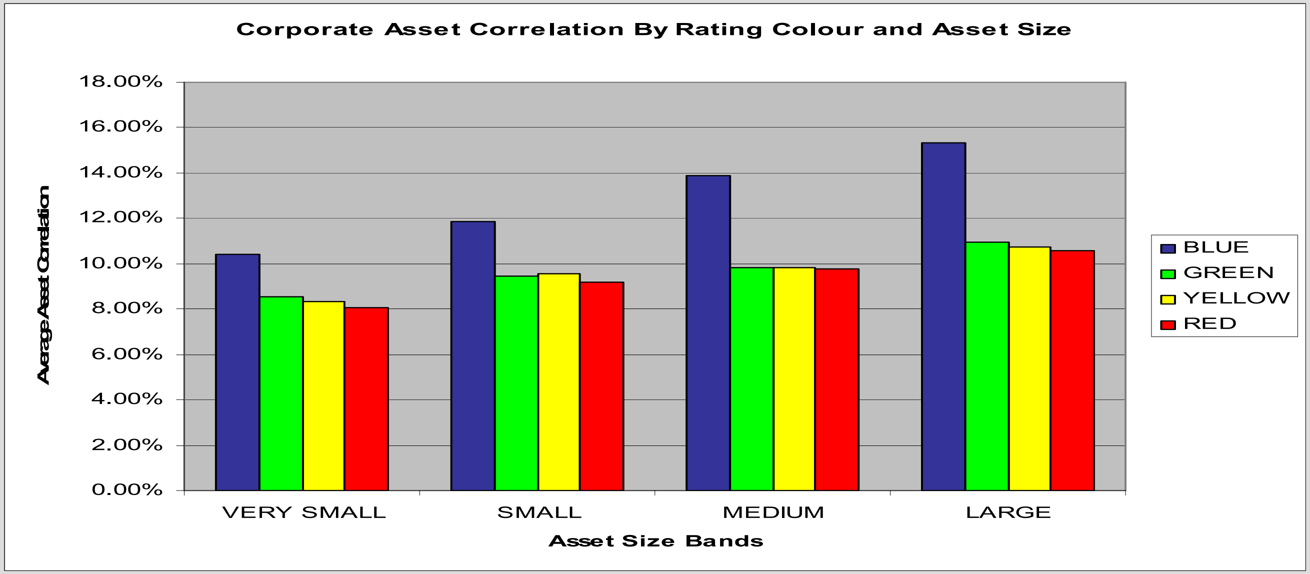
$$Corr(R_i^a, R_j^b) = \begin{cases} (\beta_a)^2 & \text{intra-cluster correlation} \\ \beta_a \beta_b \rho_{ab} & \text{inter-cluster correlation} \end{cases}$$

* The classical one factor model (as in Basle II) assumes that the factors are independent of each other. Here the multi sector factor model does not assume independent factors (as one can see in the right picture $\rho_{ab}$ has to be estimated).

# 5. Results
## 5.1. Trends

**Reality check**: We should observe the following **trends**

•Increasing asset correlations with increasing asset size
•Decreasing asset correlations with increasing default probability



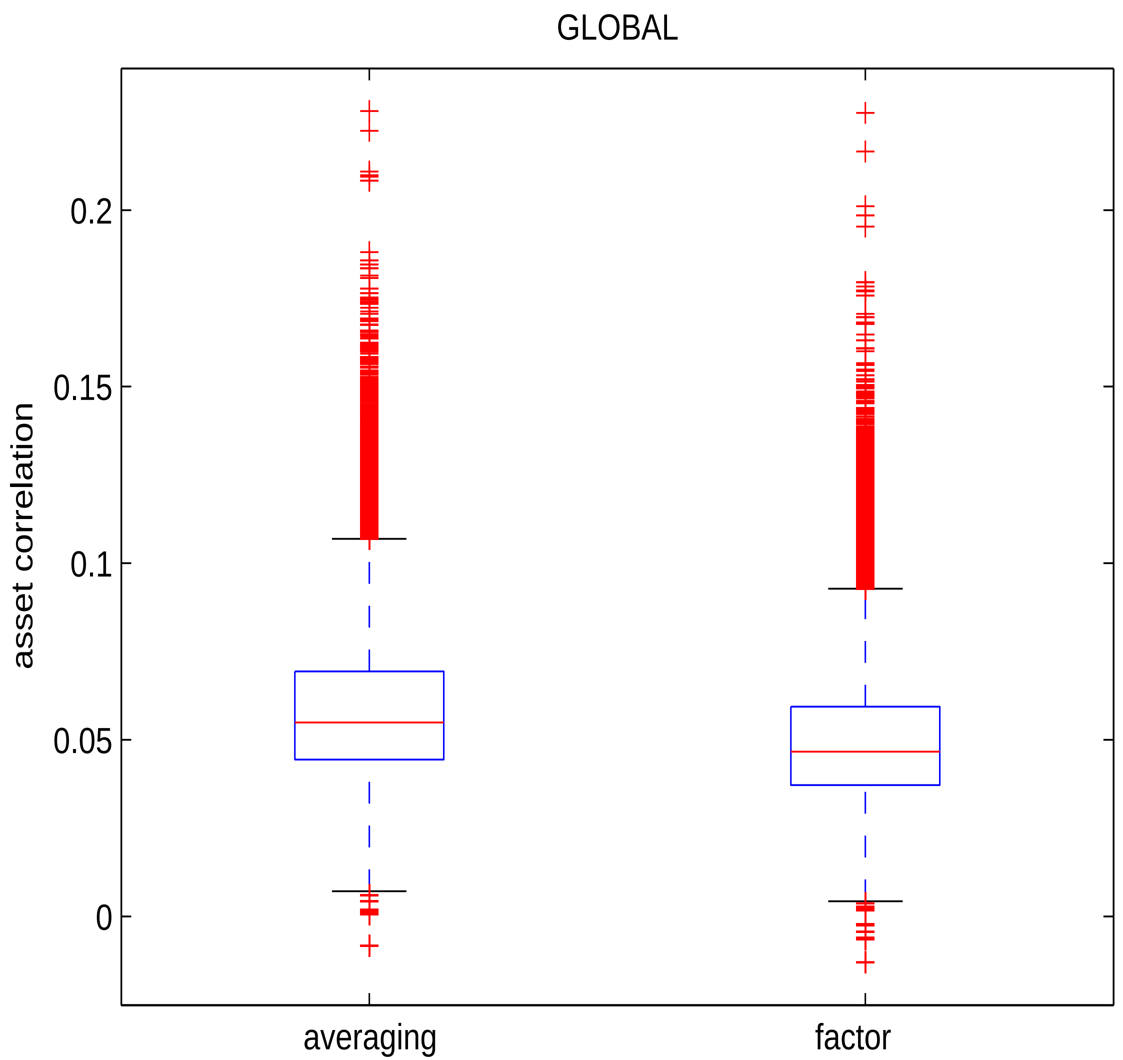
# 5.2. Comparison

| | averaging | factor |
|---|---|---|
| **mean** | 0.058 | 0.051 |
| **std** | 0.020 | 0.019 |

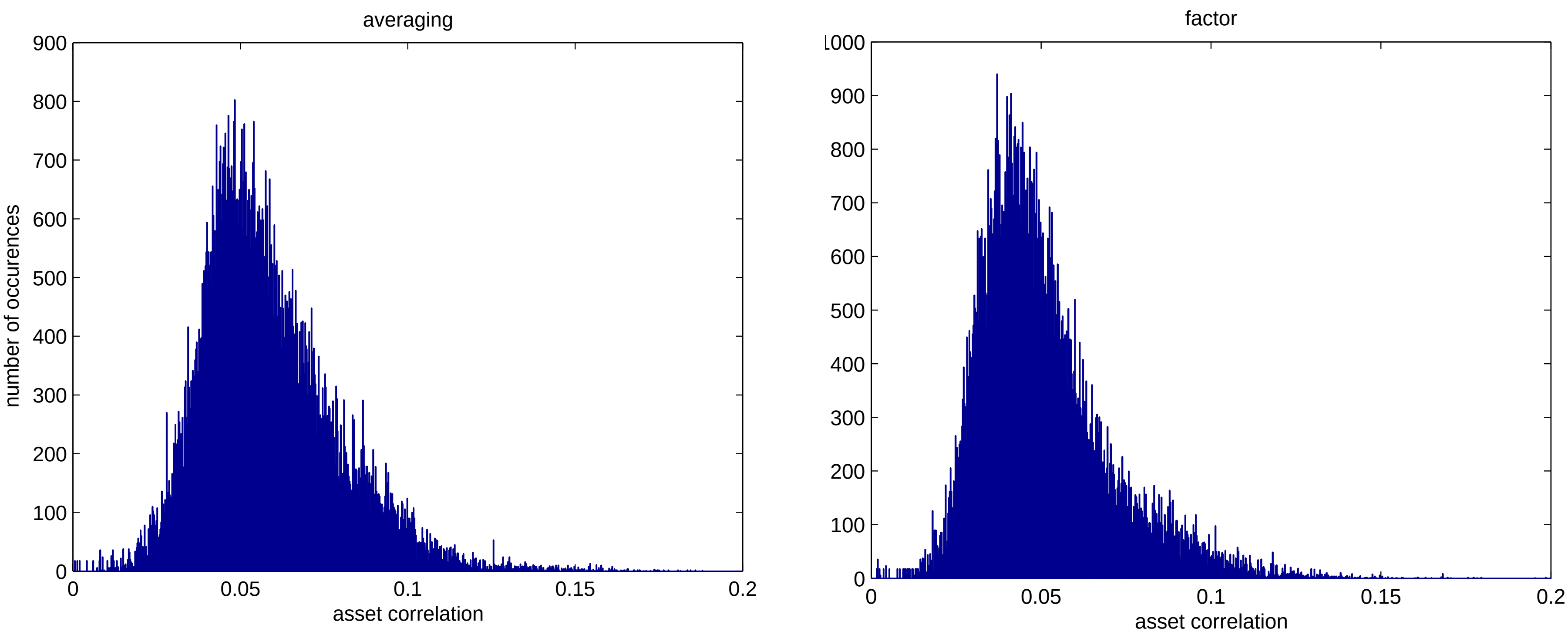
difference $\Delta \approx 1$ %

On average,
the factor model
produces slightly lower
asset correlations.

The difference is statistically significant,
but practically the results are **highly similar**.

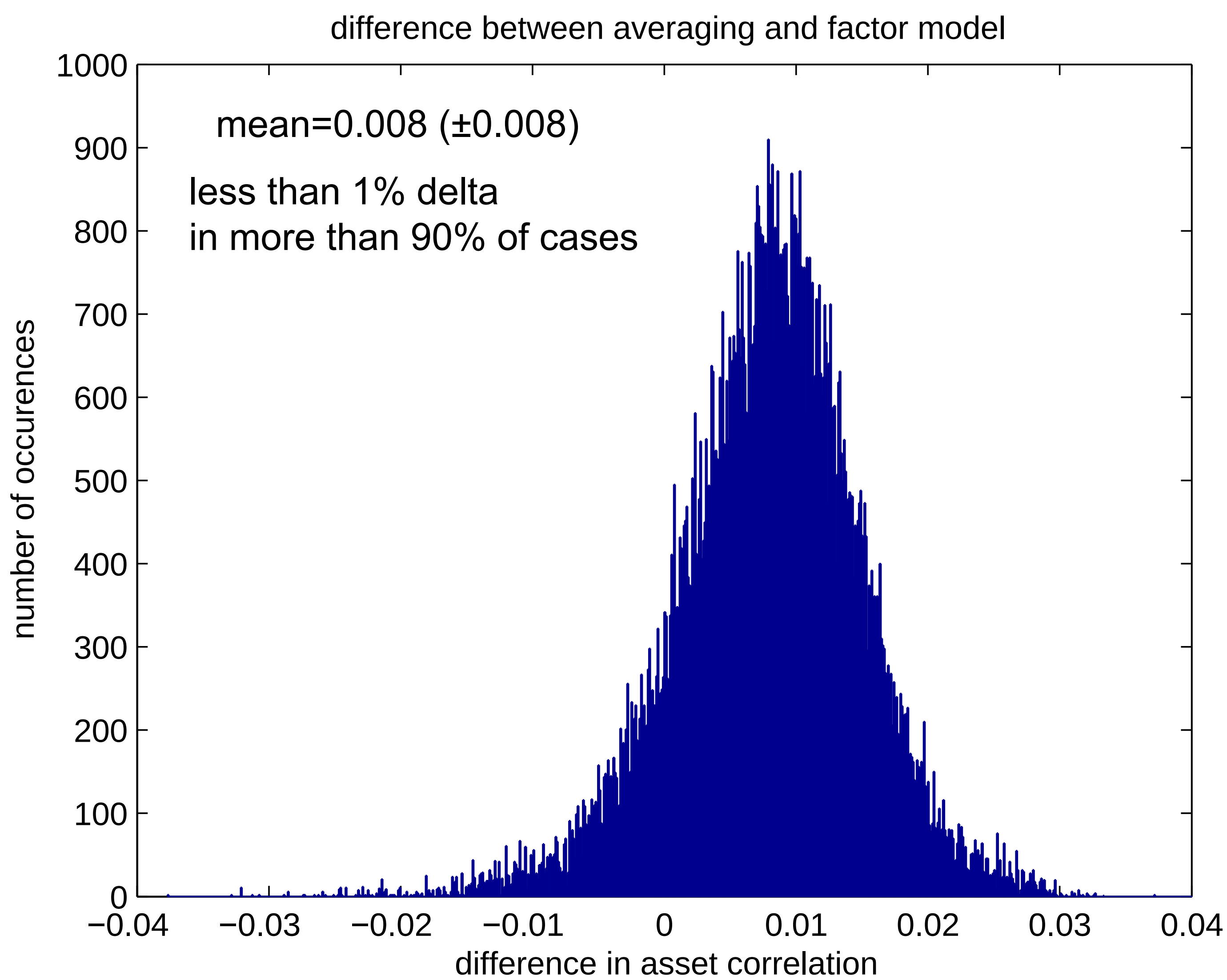


# 5.3. Histogram of overall asset correlations



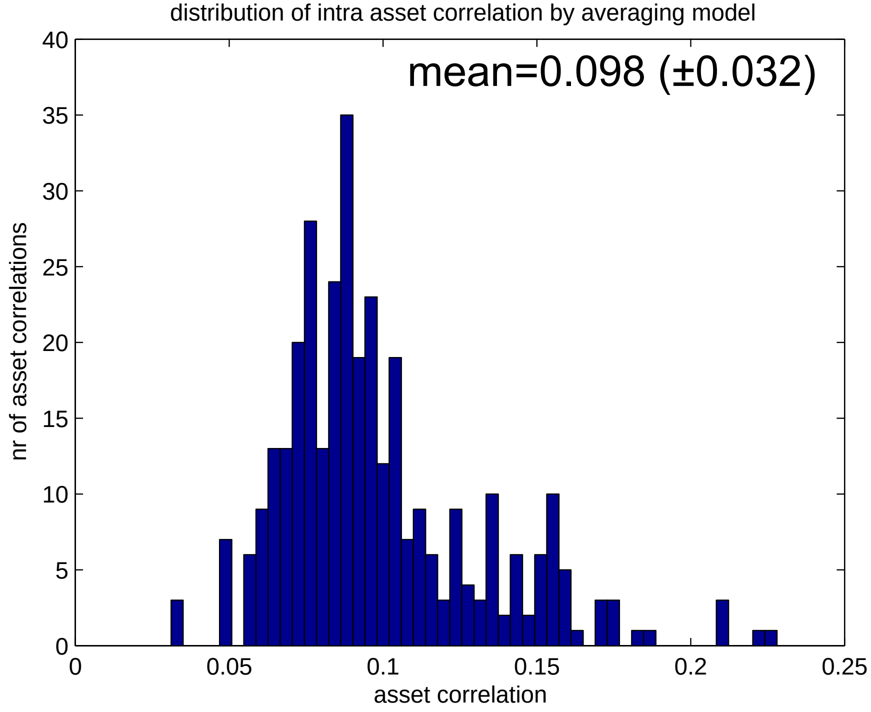

Very similar distributions, both are slightly skewed towards higher correlations.

# 5.4. Histogram of differences

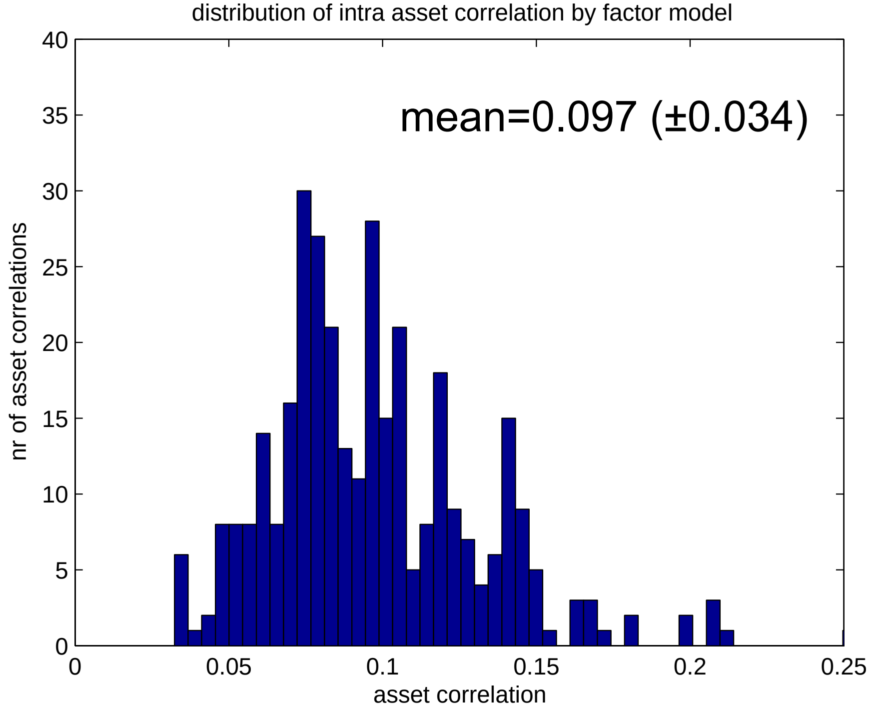


# 5.5. Histogram of intra asset correlation

averaging




factor model

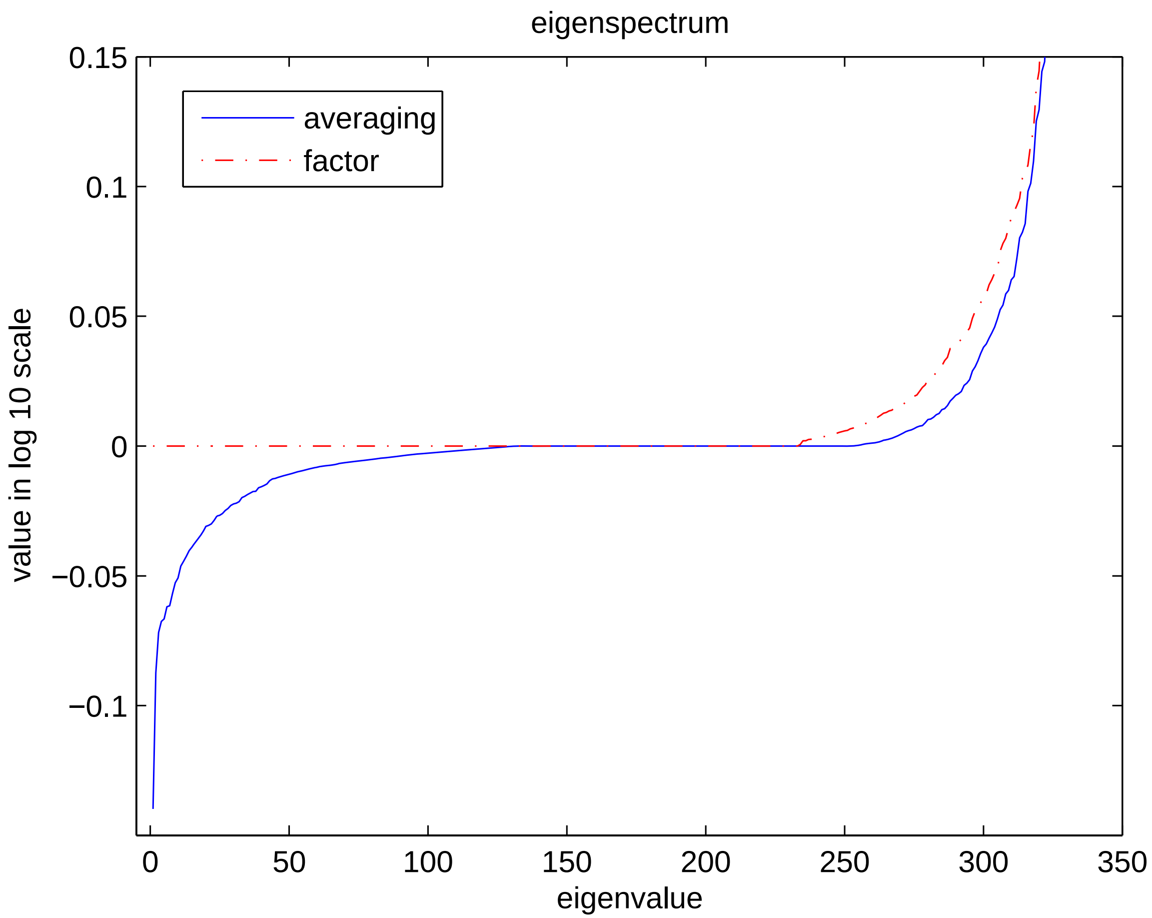



**Intra** cluster asset correlation is larger than inter correlation, but both again similar.

| | averaging | factor |
|---|---|---|
| **Corporates :** | **9.8%** | **9.7%** |
| **Individuals :** | **4.8%** | **7.4%** |
| **Publics :** | **15.4%** | **16.6%** |

FORTIS

# 5.6. Similar, but different in matrix structure




rank(averaging)= 218

rank(factor)=107

The factor model has more interdependency, induced by the single factor link.

The relative 2-norm distance is 16%, thus little structural similarity.

Averaging matrix still needs to be made positive semi definite.

FORTIS

# 5.7. Comparison with the literature

Reference: A. Chernih, S. Vanduffel, L. Henrard: Asset Correlations: Shifts and tides, 2006

| Source Study | Data Source | Results |
|---|---|---|
| Gordy (2002) | S&P | 1.5% - 12.5% |
| Cespedes (2000) | Moody's | 10% |
| Hamerle *et al.* (2003a) | | max of 2.3% |
| Hamerle *et al.* (2003b) | S&P 1982-99 | 0.4% - 6.04% |
| Frey *et al.* (2001) | UBS | 2.6%, 3.8%, 9.21% |
| Frey & McNeil (2003) | S&P 1981 - 2000 | 3.4% - 6.4% |
| Dietsch & Petey (2004) | Coface 1994-2001<br>AK 1997-2001 | 0.12% - 10.72% |
| Jobst & de Servigny (2004) | S&P 1981-2003 | intra 14.6%, inter 4.7% |
| Duellmann & Scheule (2003) | DB 1987 - 2000 | 0.5% - 6.4% |
| Jakubik (2006) | BF 1988 - 2003 | 5.7% |

Table 1: *Asset correlations from default data*

S&P: Standard and Poor's

DB: Deutsche Bundesbank

AK: Allgemeine Kredit

BF: Bank of Finland

**Interpretation**: This table clearly shows that the new asset correlation values are actually in line with the results published in the recent literature.

FORTIS

# 5.8. Are correlations too low? In fact not so

•A relatively low default correlation may cause a relatively high conditional probability

•Example Firm A: PD = 0.03 %
Firm B: PD = 2.05 %

| DC | P(B\|A) | **Interpretation** |
|---|---|---|
| 0 % | 2.05 % | Correlations do not |
| 1.50 % | 14.32 % | need to be so high |
| 11.97 % | 100 % | to cause default |

P(B|A): Probability that firm B defaults given that firm A defaults

FORTIS

# 5.9. Are correlations too low? in a sense they are

• The asset matrix was made semi positive definite (spd) by recomposing the eigenvalue decomposition where negative eigenvalues are set to zero. This procedure induces changes in the values of the matrix. This pushes the correlations somewhat up:

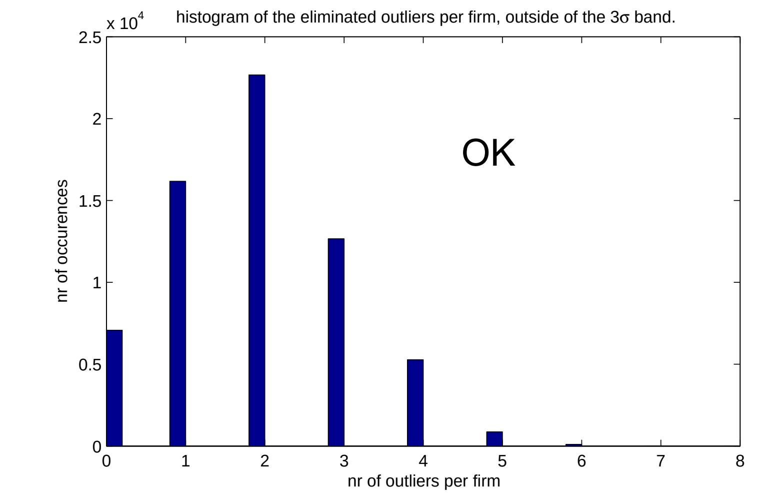
| | averaging (raw) | | averaging (spd) |
|---|---|---|---|
| **Corporates** | **9.8%** | $\Rightarrow$ | **10.4%** |
| **Individuals** | **4.8%** | $\Rightarrow$ | **6.5%** |
| **Publics** | **15.4%** | $\Rightarrow$ | **15.7%** |

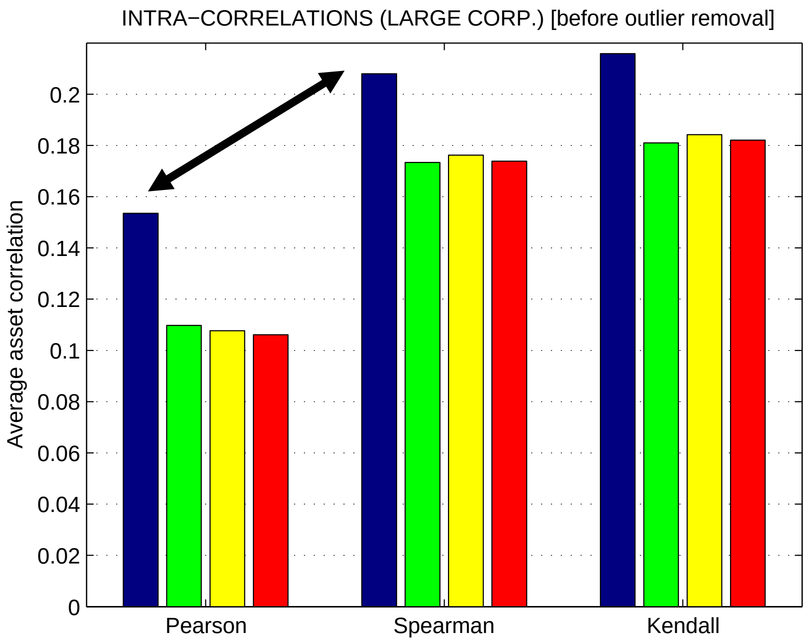
• One could acknowledge the non-normal character of asset return movement and take a more appropiate robust measure for correlation, for example rank correlations (Spearman and Kendall instead of Pearson). In case there is a large discrepancy between these, it might be necessary to consider removal of outliers by e.g. peak-shaving at $\pm 3\sigma$.

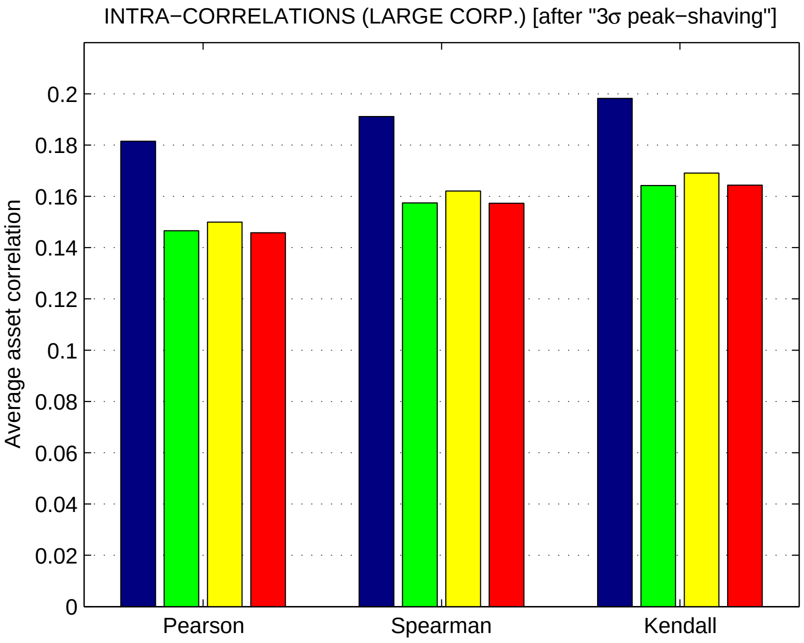
# 5.10. Peak-shaving impact

Remove of each firm all datapoints outside the band of +3σ and -3σ standard deviations around the mean. On average 2 points fall off from 107.

Impact: increases of around 30-50%, reconciliating the correlation measures.



FORTIS

# Conclusions

| METHOD | plus | minus |
|---|---|---|
| **historical data** | least biased | too few data |
| **local averaging** | less biased (semi-parametric) | 1 company, 1 cluster; intensive computations |
| **factor modelling** | 1 company, multiple clusters ; quick computation | biased by model assumptions (parametric) |

- *asset correlations are mainly in the range between 3% and 10%, or 6% to 20% if one takes into account robust correlation/peak-shaving.*

- *Inter and intra* asset correlations are very similar for both models, the single factor model yielding slightly lower correlations.