



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Overlapping Correlation Coefficient

Paolo Tasca

ETH Risk Center – Working Paper Series

ETH-RC-13-004

The **ETH Risk Center**, established at ETH Zurich (Switzerland) in 2011, aims to develop cross-disciplinary approaches to integrative risk management. The center combines competences from the natural, engineering, social, economic and political sciences. By integrating modeling and simulation efforts with empirical and experimental methods, the Center helps societies to better manage risk.

More information can be found at: <http://www.riskcenter.ethz.ch/>.

Overlapping Correlation Coefficient

Paolo Tasca

Abstract

This paper provides a mapping from portfolio risk diversification into the pairwise correlation between portfolios. In a finite market of uncorrelated assets, portfolio risk is reduced by increasing diversification. However, higher the diversification level, the greater is the overlap between portfolios. The overlap, in turn, leads to greater correlation between portfolios.

Keywords: diversification, overlapping portfolios, correlation

Classifications: *JEL Codes:* C02, G11

URL: http://web.sg.ethz.ch/ethz_risk_center_wps/ETH-RC-13-004

Notes and Comments: Status: Submitted

Overlapping Correlation Coefficient

Paolo Tasca

Chair of Systems Design, ETH Zurich, Weinbergstrasse 58, 8092 Zurich, Switzerland

Abstract

This paper provides a mapping from portfolio risk diversification into the pairwise correlation between portfolios. In a finite market of uncorrelated assets, portfolio risk is reduced by increasing diversification. However, higher the diversification level, the greater is the overlap between portfolios. The overlap, in turn, leads to greater correlation between portfolios.

JEL classification: C02 G11

1 Introduction

According to recent empirical evidence, financial sector balance sheets exhibit increasing *homogeneity*. This trend has been favored both by deregulation and by financial innovation and it is revealed in both *risk management* and *business* strategies. The upshot is the increase of return correlation across financial sectors and within them.¹ On one side, the homogenization of risk management practices reflects common *standards* and compliance with uniform regulations (see, e.g., Persaud, 2000). On the other side, the homogenization of business activities reflects common underlying *behaviors* from specialization to diversification.

I am grateful to Claudio J. Tessone, Rene Pfitzner and Frank Schweitzer. The author acknowledges financial support from the ETH Competence Center “Coping with Crises in Complex Socio-Economic Systems” CHIRP 1 (grant no. CH1-01-08-2), the European Commission FET Open Project “FOC” (grant no. 255987), and the Swiss National Science Foundation project “OTC Derivatives and Systemic Risk in Financial Networks” (grant no. CR12I1-127000/1). Correspondence to Paolo Tasca, Chair of Systems Design, ETH, Zurich, WEV G 208 Weinbergstrasse 56/58 8092 Zurich; e-mail: ptasca@ethz.ch.

¹The correlation patterns were around 0.7-0.9 when the recent 2007-2008 U.S. financial crisis reached its peak. See e.g., (Patro et al., 2012).

In this paper we draw attention to how the homogenization of investors' portfolios via *asset commonality* (i.e., common asset holding) depends on their levels of diversification.² By extending the level of portfolio diversification, investors increase the overlap between their portfolios. This, in turn, increases their correlation. The syllogism goes as follow. The higher the portfolio diversification, the greater their overlap. The higher the overlap between portfolios, the greater their pairwise correlation. The higher the correlation between portfolios, the greater the probability that investors encounter the same source of risk.

Indeed, asset commonality is widely considered to have been the primary source of the recent financial crises. In the economic literature, Stein (2009) identifies “crowding”, or similar portfolios among sophisticated investors, as a risk in financial markets. In Wagner (2008), investors with similar diversification strategies concentrated more in risky projects than liquidity may cause a liquidity shortages to become more likely as a result, and the probability of a crisis to rise unambiguously. Acharya and Yorulmazer (2004) and Acharya and Yorulmazer (2007) show that banks have incentive to herd. In so doing, they increase the risks of failing together. Also Brunnermeier and Sannikov (2010) identify portfolio overlaps as a destabilizing mechanism in financial markets. According to Allen et al. (2012) banks swap assets to diversify their individual risk. In so doing, they generate “excessive systemic risk”. (Acharya, 2009), building on the previous work of (Shaffer, 1989), explores the systemic impact that attend banks' asset-side homogeneous behavior. Elsinger et al. (2006) identify correlation in banks' asset portfolios as the main source of systemic risk. Finally, the complex systems literature highlights the potential of homogeneous portfolios to create a tension between the individually optimal and the systemically optimal diversification. (see, e.g., Arinaminpathy et al., 2012; Beale et al., 2011; May and Arinaminpathy, 2010; Nier et al., 2007)

Nevertheless, the literature on contagion due to asset commonality misses a methodological approach that maps the level of individual portfolio diversification into the degree of correlation between portfolios, via asset commonality. To this end, we introduce the *overlapping correlation coefficient* (hereafter referred to as OCC) as a measure of the linear correlation (dependence) between two portfolios based on their level of diversification. The OCC between the portfolios P_i and P_j , randomly composed of n_i and n_j assets, is denoted as $E[Cor(P_i, P_j)]$. It gives the expected correlation between P_i and P_j , *without* any information on individual portfolio allocations but for their levels of diversification n_i and n_j , and the market size N . It assumes values between 0 and +1 inclusive, where 1 is total positive correlation, 0 is no correlation.

²While, we consider homogenization as driven by diversification, there can also be homogenization driven by herding.

To isolate the contribution of diversification to the value of OCC, the assets are assumed to be independent and identically distributed and uncorrelated. Notice that the assumption of assets being uncorrelated is not a simplifying device adopted for analytical convenience. Rather, it allows to understand how the OCC increases solely with the contribution of the asset commonality via portfolio diversification. Then, bringing into the analysis an arbitrary correlation matrix between assets would mislead the concept and alter the results of the OCC. It must be finally considered that in reality the majority of the assets are positively correlated with the market, and only very few of them may have negative correlation. Therefore, the OCC represents a *lower bound* of the real correlation coefficient between portfolios.

To conclude, the OCC may have general applications beyond the financial context. More broadly, the OCC captures the correlation between unexpected outcomes that individual strategies, independently chosen from a finite set of mutually exclusive alternatives, may generate.

2 Portfolio Construction

Consider a frictionless finite market composed of a N -set $\{X_1, \dots, X_N\}$ of risky assets and a group of M investors. We assume that the assets are *indistinguishable* and *uncorrelated*. This means that $E(X_\ell)_{t \geq 0} = \mu$ and $\text{Var}(X_\ell)_{t \geq 0} = \sigma^2$ for all $\ell = 1, \dots, N$ and $\text{Cov}(X_\ell, X_y)_{t \geq 0} = 0$ for all $\ell \neq y$. To keep the notations simple, in the following we omit the time sub-index t .

In a fine-grained portfolio each asset is equally-weighted and represents only a small fraction of the total portfolio value. The idiosyncratic risks carries by the assets can therefore be mitigated by holding such a portfolio. Since assets are indistinguishable and uncorrelated, the portfolio $P_i = \left(\frac{1}{n_i}X_1 + \dots + \frac{1}{n_i}X_N\right)$ composed of an equally-weighted linear combination of $n_i \leq N$ assets chosen among the N assets available for investment is also the optimal allocation strategy. n_i defines the level of diversification of P_i . The portfolio variance asymptotically decreases in n_i :

$$\begin{aligned}
 \text{Var}(P_i) &= \text{Var}\left(\frac{1}{n_i}X_1 + \dots + \frac{1}{n_i}X_N\right) \\
 &= \frac{1}{n_i^2}\text{Var}(X_1) + \dots + \frac{1}{n_i^2}\text{Var}(X_N) \\
 &= \frac{1}{n_i^2} \sum_{\ell=1}^N \text{Var}(X_\ell) = \frac{1}{n_i^2} \sum_{\ell=1}^N \sigma^2 \\
 &= \frac{\sigma^2}{n_i}.
 \end{aligned} \tag{1}$$

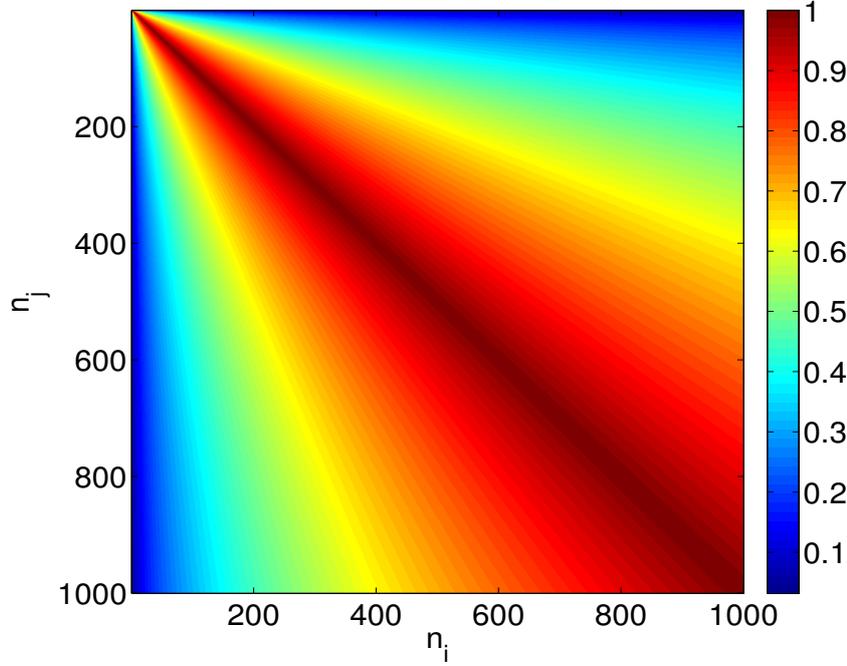


Figure 1: Upper-bound of the correlation between two portfolios P_i and P_j with respect to their specific level of diversification. Market size $N = 1000$.

The covariance between any portfolio P_i and P_j held by investors i and j respectively, is:

$$\begin{aligned} Cov(P_i, P_j) &= Cov\left(\frac{1}{n_i}X_1 + \dots + \frac{1}{n_i}X_N, \frac{1}{n_j}X_1 + \dots + \frac{1}{n_j}X_N\right) \\ &= \frac{1}{n_i} \frac{1}{n_j} \sum_{\ell=1}^N \sum_{y=1}^N Cov(X_\ell, X_y). \end{aligned} \quad (2)$$

Then, the Pearson correlation between P_i and P_j is:

$$Cor(P_i, P_j) = \frac{1}{n_i n_j} \sum_{\ell=1}^N \sum_{y=1}^N Cov(X_\ell, X_y) / \left(\frac{\sigma^2}{\sqrt{n_i n_j}} \right). \quad (3)$$

The correlation coefficient has an upper bound at $\frac{\min(n_i, n_j) \sigma^2}{n_i n_j} \times \frac{\sqrt{n_i n_j}}{\sigma^2}$ that is equivalent to $\frac{\min(n_i, n_j)}{\sqrt{n_i n_j}}$. In a hypothetical market with infinite size, viz. $N \rightarrow \infty$, the upper bound is equal to one when

both n_i and n_j tend to N and it is equal to zero when only n_i (or n_j) tends to N . To graphically understand how the upper-bound of the correlation changes with respect to the levels of diversification, Figure 1 shows the result of a numerical example for a market with a size $N = 1000$.

3 Overlapping Correlation

In order to formulate the overlapping correlation coefficient (OCC) let us start by considering that assets are uncorrelated and indistinguishable. Since assets are uncorrelated, the only source of correlation between any portfolios P_i and P_j comes from *asset commonality*, i.e. the level of their overlap. Since assets are indistinguishable investors are neutral w.r.t. any asset X_ℓ and X_y for all $\ell \neq y$ in the N -set. More formally, let $C(N, n_i)$ be the set of n_i -subsets of the N -set. The set $C(N, n_i)$ has cardinality ${}_N C_{n_i}$ and contains all the possible portfolios P_i^1, P_i^2, \dots with size n_i . Since assets are indistinguishable we use the convention $P_i^a \sim P_i^b$ to indicate that the portfolio “ P_i^a is indifferent to P_i^b ” for all $P_i^a, P_i^b \in C(N, n_i)$. In other terms, each investor i is indifferent between any possible portfolio chosen from ${}_N C_{n_i}$ combinations of n_i assets taken from the N -set of assets available for investment.

In the case $n_i \geq n_j$, the general expression for the OCC is³:

$$E[Cor(P_i, P_j)] = \frac{{}_{N-n_i} C_{N-n_j} {}_{n_i} C_{n_i}}{{}_N C_{n_j}} \left(\frac{n_i}{\sqrt{n_i n_j}} \right) + \frac{{}_{N-n_i} C_{n_j-n_i+l} {}_{n_i} C_{n_i-l}}{{}_N C_{n_j}} \left(\frac{n_i-l}{\sqrt{n_i n_j}} \right) \quad (4)$$

$$\equiv \frac{\sqrt{n_i n_j}}{N} \quad (5)$$

where $l = 1, \dots, n_i$ and with the condition that $l \leq N - n_j$. Figure 2 shows how the OCC monotonously increases with n_i and n_j from 0 to 1.

As an illustration to understand our combinatorial approach used to derived the OCC as a function of the tuple (N, n_i, n_j) , let us consider a 5-asset market $\{X_1, \dots, X_5\}$. Then, among the five assets available for investment, let investor i equally diversify across three randomly selected assets, i.e., $n_i = 3$. Then, $C(5, 3)$ is the set of ${}_5 C_3 = 10$ possible portfolios composed of three assets each: $P_i^1 = \{1, 2, 3\} \sim P_i^2 = \{1, 2, 4\} \sim P_i^3 = \{1, 2, 5\} \sim P_i^4 = \{1, 3, 4\} \sim P_i^5 = \{1, 3, 5\} \sim P_i^6 = \{1, 4, 5\} \sim P_i^7 = \{2, 3, 4\} \sim P_i^8 = \{2, 3, 5\} \sim P_i^9 = \{2, 4, 5\} \sim P_i^{10} = \{3, 4, 5\}$. Similarly, let j equally diversify across four randomly selected assets, i.e., $n_j = 4$. Then, $C(5, 4)$ is the set of ${}_5 C_4 = 5$ possible portfolios composed of four assets each: $P_j^1 = \{1, 2, 3, 4\} \sim$

³Since the correlation matrix is symmetric, by inverting the sub-indexes i with j one can obtain the expected correlation in the case $n_i \geq n_j$.

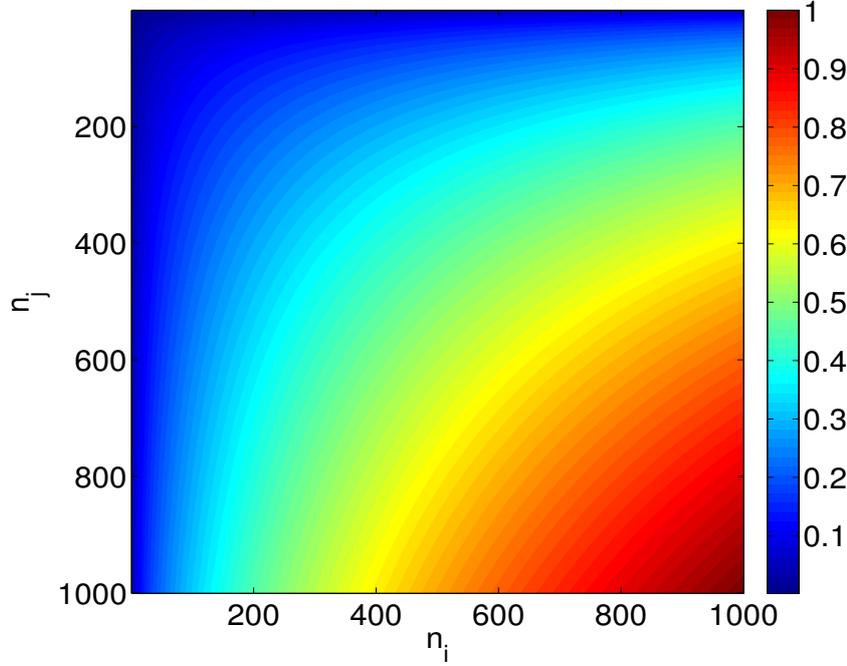


Figure 2: Correlation surface between P_i and P_j for different diversification levels n_i and n_j . Market size $N = 1000$.

$P_j^2 = \{1, 2, 3, 5\} \sim P_j^3 = \{1, 3, 4, 5\} \sim P_j^4 = \{2, 3, 4, 5\} \sim P_j^5 = \{1, 2, 4, 5\}$. For any portfolio $P_i^a \in C(5, 3)$ there are two portfolios $P_j^a, P_j^b \in C(5, 4)$ overlapping for $3/4$ with P_i^a and three portfolios $P_j^a, P_j^b, P_j^c \in C(5, 4)$ overlapping for $2/5$ with P_i^a . See Table 1. Therefore, in our example, for any portfolio P_i^a chosen by i , with probability $p = \frac{3}{5}$ investor j might chose a portfolio P_j^a that overlaps for $2/5$ with P_i^a and with probability $1 - p = \frac{2}{5}$, investor j might chose a portfolio P_j^a that overlaps for $3/4$ with P_i^a .⁴

If i chooses the portfolio P_i^1 , then P_j^1 and P_j^2 overlap for $3/4$ with P_i^1 . Instead, P_j^3, P_j^4 and P_j^5 overlap for $2/5$ with P_i^1 . The pairs of portfolios with the same overlap have also the same level of correlation:

$$\begin{aligned} \text{Cor}(P_i^1, P_j^1) &= \text{Cov}\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \frac{1}{4}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + \frac{1}{4}X_4\right) / \sqrt{\text{Var}(P_i^1)\text{Var}(P_j^1)} \\ &= \frac{\frac{1}{3}\frac{1}{4}\text{Var}(X_1) + \frac{1}{3}\frac{1}{4}\text{Var}(X_2) + \frac{1}{3}\frac{1}{4}\text{Var}(X_3)}{\sqrt{\left(\frac{1}{3^2}\text{Var}(X_1) + \frac{1}{3^2}\text{Var}(X_2) + \frac{1}{3^2}\text{Var}(X_3)\right)\left(\frac{1}{4^2}\text{Var}(X_1) + \frac{1}{4^2}\text{Var}(X_2) + \frac{1}{4^2}\text{Var}(X_3) + \frac{1}{4^2}\text{Var}(X_4)\right)}} \end{aligned}$$

⁴Notice that we used the Jaccard index to measure the level of overlap.

| Portfolio P_i | Overlap between P_i and P_j for a give choice of P_i | | | | |
|--------------------------|--|--------------------------|--------------------------|--------------------------|--------------------------|
| | overlap with $P_j = 3/4$ | | overlap with $P_j = 2/5$ | | |
| $P_i^1 = \{1, 2, 3\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |
| $P_i^2 = \{1, 2, 4\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^5 = \{1, 2, 4, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ |
| $P_i^3 = \{1, 2, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ |
| $P_i^4 = \{1, 3, 4\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |
| $P_i^5 = \{1, 3, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |
| $P_i^6 = \{1, 4, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ |
| $P_i^7 = \{2, 3, 4\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |
| $P_i^8 = \{2, 3, 5\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |
| $P_i^9 = \{2, 4, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ |
| $P_i^{10} = \{3, 4, 5\}$ | $P_j^3 = \{1, 3, 4, 5\}$ | $P_j^4 = \{2, 3, 4, 5\}$ | $P_j^1 = \{1, 2, 3, 4\}$ | $P_j^2 = \{1, 2, 3, 5\}$ | $P_j^5 = \{1, 2, 4, 5\}$ |

Table 1: Possible levels of overlap between P_i and P_j .

$$= \frac{\frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2}{\frac{\sigma^2}{\sqrt{3} \sqrt{4}}} = \frac{3}{\sqrt{3} \sqrt{4}}.$$

$$\begin{aligned} \text{Cor}(P_i^1, P_j^2) &= \text{Cov}\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \frac{1}{4}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + \frac{1}{4}X_5\right) / \sqrt{\text{Var}(P_i^1)\text{Var}(P_j^2)} \\ &= \frac{\frac{1}{3} \frac{1}{4} \text{Var}(X_1) + \frac{1}{3} \frac{1}{4} \text{Var}(X_2) + \frac{1}{3} \frac{1}{4} \text{Var}(X_3)}{\sqrt{\left(\frac{1}{32} \text{Var}(X_1) + \frac{1}{32} \text{Var}(X_2) + \frac{1}{32} \text{Var}(X_3)\right) \left(\frac{1}{42} \text{Var}(X_1) + \frac{1}{42} \text{Var}(X_2) + \frac{1}{42} \text{Var}(X_3) + \frac{1}{42} \text{Var}(X_5)\right)}} \\ &= \frac{\frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2}{\frac{\sigma^2}{\sqrt{3} \sqrt{4}}} = \frac{3}{\sqrt{3} \sqrt{4}}. \end{aligned}$$

$$\begin{aligned} \text{Cor}(P_i^1, P_j^3) &= \text{Cov}\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \frac{1}{4}X_1 + \frac{1}{4}X_3 + \frac{1}{4}X_4 + \frac{1}{4}X_5\right) / \sqrt{\text{Var}(P_i^1)\text{Var}(P_j^3)} \\ &= \frac{\frac{1}{3} \frac{1}{4} \text{Var}(X_1) + \frac{1}{3} \frac{1}{4} \text{Var}(X_3)}{\sqrt{\left(\frac{1}{32} \text{Var}(X_1) + \frac{1}{32} \text{Var}(X_2) + \frac{1}{32} \text{Var}(X_3)\right) \left(\frac{1}{42} \text{Var}(X_1) + \frac{1}{42} \text{Var}(X_3) + \frac{1}{42} \text{Var}(X_4) + \frac{1}{42} \text{Var}(X_5)\right)}} \\ &= \frac{\frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2}{\frac{\sigma^2}{\sqrt{3} \sqrt{4}}} = \frac{2}{\sqrt{3} \sqrt{4}}. \end{aligned}$$

$$\begin{aligned} \text{Cor}(P_i^1, P_j^4) &= \text{Cov}\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \frac{1}{4}X_2 + \frac{1}{4}X_3 + \frac{1}{4}X_4 + \frac{1}{4}X_5\right) / \sqrt{\text{Var}(P_i^1)\text{Var}(P_j^4)} \\ &= \frac{\frac{1}{3} \frac{1}{4} \text{Var}(X_2) + \frac{1}{3} \frac{1}{4} \text{Var}(X_3)}{\sqrt{\left(\frac{1}{32} \text{Var}(X_1) + \frac{1}{32} \text{Var}(X_2) + \frac{1}{32} \text{Var}(X_3)\right) \left(\frac{1}{42} \text{Var}(X_2) + \frac{1}{42} \text{Var}(X_3) + \frac{1}{42} \text{Var}(X_4) + \frac{1}{42} \text{Var}(X_5)\right)}} \\ &= \frac{\frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2}{\frac{\sigma^2}{\sqrt{3} \sqrt{4}}} = \frac{2}{\sqrt{3} \sqrt{4}}. \end{aligned}$$

$$\text{Cor}(P_i^1, P_j^5) = \text{Cov}\left(\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \frac{1}{4}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_4 + \frac{1}{4}X_5\right) / \sqrt{\text{Var}(P_i^1)\text{Var}(P_j^5)}$$

$$\begin{aligned}
&= \frac{\frac{1}{3} \frac{1}{4} \text{Var}(X_1) + \frac{1}{3} \frac{1}{4} \text{Var}(X_2)}{\sqrt{\left(\frac{1}{3^2} \text{Var}(X_1) + \frac{1}{3^2} \text{Var}(X_2) + \frac{1}{3^2} \text{Var}(X_3)\right) \left(\frac{1}{4^2} \text{Var}(X_1) + \frac{1}{4^2} \text{Var}(X_2) + \frac{1}{4^2} \text{Var}(X_4) + \frac{1}{4^2} \text{Var}(X_5)\right)}} \\
&= \frac{\frac{1}{3} \frac{1}{4} \sigma^2 + \frac{1}{3} \frac{1}{4} \sigma^2}{\frac{\sigma^2}{\sqrt{3} \sqrt{4}}} = \frac{2}{\sqrt{3} \sqrt{4}}.
\end{aligned}$$

Then, $\text{Cor}(P_i^1, P_j^1) = \text{Cor}(P_i^1, P_j^2) = \frac{3}{\sqrt{3} \sqrt{4}}$ and $\text{Cor}(P_i^1, P_j^3) = \text{Cor}(P_i^1, P_j^4) = \text{Cor}(P_i^1, P_j^5) = \frac{2}{\sqrt{3} \sqrt{4}}$.

Since, it is not known, *a priori*, which portfolios P_i^a and P_j^a will be chosen from their respective set, the overlapping correlation between any randomly selected portfolio $P_i^a \in C(5, 3)$ composed of three projects out of five and any randomly selected portfolio $P_j^a \in C(5, 4)$ composed of four projects out of five, is given in probabilistic terms as follows:

$$\begin{aligned}
E[\text{Cor}(P_i, P_j)] &= p \times \text{Cor}(P_i^1, P_j^1) + (1 - p) \times \text{Cor}(P_i^1, P_j^3) & (6) \\
&= p \times \text{Cor}(P_i^1, P_j^1) + (1 - p) \times \text{Cor}(P_i^1, P_j^4) \\
&\equiv p \times \text{Cor}(P_i^1, P_j^1) + (1 - p) \times \text{Cor}(P_i^1, P_j^5) \\
&\equiv p \times \text{Cor}(P_i^1, P_j^2) + (1 - p) \times \text{Cor}(P_i^1, P_j^3) \\
&\equiv p \times \text{Cor}(P_i^1, P_j^2) + (1 - p) \times \text{Cor}(P_i^1, P_j^4) \\
&\equiv p \times \text{Cor}(P_i^1, P_j^2) + (1 - p) \times \text{Cor}(P_i^1, P_j^5) \\
&= \frac{3}{5} \times \frac{2}{\sqrt{3} \sqrt{4}} + \frac{2}{5} \times \frac{3}{\sqrt{3} \sqrt{4}} \\
&= \frac{1}{5} \left(\frac{3 \times 4}{\sqrt{3} \sqrt{4}} \right) \\
&= \frac{\sqrt{3} \times 4}{5} & (7)
\end{aligned}$$

In the general case, Eq. (7) can be written as $E[\text{Cor}(P_i, P_j)] = \frac{\sqrt{n_i n_j}}{N}$, which is equivalent to Eq. (5).

4 Discussion

The correlation coefficient $\text{Cor}(P_i, P_j)$ in Eq. (3) depends on the market size N , on the levels of diversification n_i, n_j and on the specific assets $X_{\ell=1, \dots, N}$ chosen by both i and j . Therefore, $\text{Cor}(P_i, P_j)$ can be considered a *backward looking* quantity that depends on the specific portfolio allocation adopted by both i and j . On the contrary, the OCC is a *forward looking* quantity that depends only on: (1) the number N of assets available for investment; (2) the number $n_i \leq N$

and $n_j \leq N$ of assets held in P_i and P_j , respectively; (3) the cardinality ${}_N C_{n_i}$ and ${}_N C_{n_j}$ of the sets $C(N, n_i)$ and $C(N, n_j)$. Therefore, to compute the OCC we do *not* need to know which assets exactly compose P_i and P_j . We simply need to know their levels of diversification, i.e., n_i and n_j .

5 Conclusions

In the current financial arena where the diversity of investors has been gradually eroded by the sharing of common assets, the study of the exact functional relation between the extension of risk diversification and the portfolio correlation between investors (via common asset holding), is of dramatic importance. This is especially true if the risk management strategy of full diversification is efficient from a single-investor point of view and inefficient from a system perspective. The contribution of the paper is to present an analytical solution to this problem by mapping the levels of portfolio diversification into their pairwise correlation.

References

- Acharya, V. (2009). A theory of systemic risk and design of prudential bank regulation. *Journal of Financial Stability*, 5(3):224–255.
- Acharya, V. and Yorulmazer, T. (2004). Limited liability and bank herding.
- Acharya, V. and Yorulmazer, T. (2007). Too many to fail. an analysis of time-inconsistency in bank closure policies. *Journal of Financial Intermediation*, 16(1):1–31.
- Allen, F., Babus, A., and Carletti, E. (2012). Asset commonality, debt maturity and systemic risk. *Journal of Financial Economics*, 104(3):519–534.
- Arinaminpathy, N., Kapadia, S., and May, R. (2012). Size and complexity in model financial systems. *Proceedings of the National Academy of Sciences*, 109(45):18338–18343.
- Beale, N., Rand, D., Battey, H., Croxson, K., May, R., and Nowak, M. (2011). Individual versus systemic risk and the regulator’s dilemma. *Proceedings of the National Academy of Sciences*, 108(31):12647–12652.
- Brunnermeier, M. and Sannikov, Y. (2010). A macroeconomic model with a financial sector. *Unpublished manuscript, Princeton University*.

- Elsinger, H., Lehar, A., and Summer, M. (2006). Risk Assessment for Banking Systems. *Management Science*, 52(9):1301–1314.
- May, R. and Arinaminpathy, N. (2010). Systemic risk: the dynamics of model banking systems. *Journal of the Royal Society Interface*, 7(46):823–838.
- Nier, E., Yang, J., Yorulmazer, T., and Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control*, 31(6):2033–2060.
- Patro, D., Qi, M., and Sun, X. (2012). A simple indicator of systemic risk. *Journal of Financial Stability*.
- Persaud, A. (2000). Sending the herd off the cliff edge: the disturbing interaction between herding and market-sensitive risk management practices. *Journal of Risk Finance, The*, 2(1):59–65.
- Shaffer, S. (1989). Pooling intensifies joint failure risk.
- Stein, J. (2009). Presidential address: Sophisticated investors and market efficiency. *The Journal of Finance*, 64(4):1517–1548.
- Wagner, W. (2008). The homogenization of the financial system and financial crises. *Journal of Financial Intermediation*, 17(3):330–356.