# Observed Correlations and Dependencies Among Operational Losses in the ORX Consortium Database

Eric Cope and Gianluca Antonini
IBM Zurich Research Lab
Säumerstrasse 4
CH-8803 Rüschlikon
Switzerland
{erc,gan}@zurich.ibm.com
+41 44 724 8423

in collaboration with
The ORX Analytics Working Group*

November 27, 2008

*Significant contributions of the ORX Analytics Working Group are acknowledged, in particular those of John Walter, Giulio Mignola, Mark Piche, Roberto Ugoccioni, Meirong Li, Rocco Quartu, and Peter Schaller.

**Abstract**

We survey a range of correlation and dependence measures among operational losses among an international consortium (ORX) of banks. In general, we find little evidence of strong correlations, and some slight evidence of tail dependencies among quarterly aggregate loss values among business line, event type, and Basel cell (combination business line/event type) units of measure. The implications for diversification benefits when aggregating losses across units of measure are further explored through direct empirical measurement.

Keywords: Operational losses, correlation, tail dependence, consortium data, diversification benefit

# 1 Introduction

Under the Advanced Measurement Approach (AMA) outlined in the Basel II Accord, banks must set aside regulatory capital equal to an estimate of the size of the total losses sustained within a year that would occur with 0.1% likelihood. Because a bank's total losses are made up of contributions from a heterogeneous group of loss event types that affect various lines of business, estimates of total losses are built from loss models for these various loss subcategories. Typically, this is done using the Loss Distribution Approach (LDA), which computes loss distribution estimates based on estimates of the frequency and severity of losses in each of several categories, and then combines these category-level distributional estimates into an estimate of the total, bank-wide loss distribution. Banks are allowed the option to incorporate internal estimates of dependence within this framework, if they can demonstrate the soundness, robustness, and integrity of their internal estimates (Basel Committee for Banking Supervision 2003; Dugan, Johnson, Feldman, and Reich 2007).

Gaining an accurate understanding of the degrees of dependence among losses in various units of measure is a critical factor in estimating the total capital charge under the LDA approach. On the one hand, if perfect correlations exist among losses in different business lines, then if an extreme loss event occurs in one business line, such an extreme event will also be realized in the other business lines as well. On the other hand, if losses are realized independently, then there is a potential for sizable "diversification benefits" as large losses need not occur simultaneously in various units of measure. However, several authors have pointed out that diversification benefits are not always guaranteed when losses are not perfectly correlated, even when they are independent; in fact, the once-per-thousand year total loss event can even be *higher* in cases where the losses in various units of measure are independent as when they are perfectly correlated. This apparently paradoxical result is a manifestation of the fact that the quantile function (sometimes referred to as Value at Risk) is not a coherent risk measure (Artzner, Delbaen, Eber, and Heath 1999). The effect can occur in the presence of heavy-tailed loss data, as we observe in operational loss data, even when the losses in each component category are independent of each other.

In this article, we shall report on measured correlations and dependencies among various units of measure as reported by the member banks of the ORX consortium. The Operational Riskdata eXchange (ORX) is an international consortium of banks whose primary mission is to enable the collection and mutual exchange of operational loss data. With loss records dating back to 2002, ORX currently has the most comprehensive cross-firm loss database for the banking industry, with more than 90,000 loss records collected from 41 banks, which range widely in size, line of business concentration, and region. Our goal in this study was to understand correlation and dependence behavior across the consortium to help the member banks to better meet the regulatory demands regarding correlation data, as well as to investigate the effects of dependence with respect to the LDA framework. Our main findings can be summarized in the following points:

- With few exceptions, Kendall rank correlations among quarterly aggregate losses at the business line or event type level are low, usually not exceeding 0.2.
- We generally find homogeneity among correlations measured at different banks and the correlation averages across banks, indicating that an average correlation matrix may serve as a useful reference for most consortium members.
- In the presence of heavy-tailed data, a more important measure of association than correlation may be that of tail dependence. We generally find slight evidence that extreme losses in one unit of measure are much more likely to occur when extreme losses are observed in other units of measure.
- Based on the available data, there are some indications of diversification benefits at high quantile levels of the quarterly loss distribution, although no accurate estimates of this benefit at the 99.9th percentile level can be drawn.

The paper is organized as follows. Section 2 details the data used in the study. Section 3 then describes the scope of the current analysis, against the background of past studies of correlation and dependence. Next, in Section 4, we discuss the results of the correlation analysis, focusing on the Kendall correlations of quarterly aggregate loss data at the business line, event type, and cell level. Section 5 then reports on tail probability measures and the numbers of simultaneous exceedances of high quantile levels observed across many different units of measure within the same time period. We then present some empirical findings on diversification benefits through measurements of the additivity of the loss distributions at various quantile levels in Section 6. We provide in Section 7 a discussion of the factors contributing to correlations, including long-term patterns of growth and decline in losses, the effects of exchange rate variation, and the implications of correlations based on quarterly loss data on corresponding annual losses. The conclusions of the study are finally summarized in Section 8.

## 2 Data Used in the Study

The Operational Riskdata eXchange Association (ORX) is the world's leading operational risk loss data consortium for the financial services industry. Its members use the ORX database for statistical modelling, benchmarking of loss performance, and validation of internal data collection. As a leading industry group, ORX additionally provides a forum for discussion among banks on operational risk issues and the development of industry standards. The association is further committed to advancing fundamental research into operational risk on the basis of its loss database.

As per the ORX loss reporting standards (Operational Riskdata eXchange Association 2007), loss data is submitted by member institutions to the consortium's data custodian, and is subjected to a quality control process that enforces the completeness, accuracy, and consistency of the data. Each loss in the database is categorized according to primary and secondary business lines and event types; the primary categories, which shall be our focus in this study, are listed in Table 2. Note that this categorization differs slightly from the Basel II loss taxonomy, although it is straightforward to map the categories of the two classification schemes to each other. The database records the amounts of the gross loss and of any direct and indirect recovery, the dates of occurrence, discovery, and recognition, the country in which the loss was incurred, and an indicator of whether the loss is related to credit or market risk events.

The study analyzed losses that were recognized by the member banks during the period spanning January 1, 2002 to December 31, 2007. The date of recognition of the loss was used as it is generally considered to be more accurate than either the dates of occurrence or discovery, as it is the reference date in regulatory loss reporting. It should be noted, however, that the recognition process may introduce certain biases to the timing of the losses, if losses are aggregated in time periods that are too granular. For example, since losses tend to be recognized more frequently at the ends of quarters, monthly aggregate losses may show increased correlations due to the skewness in timing. Although there do appear to be slight biases toward reporting losses at the ends of years, as we shall see in the results below, correlations at the quarterly level typically do not appear very strong. There are, however, some portions of this study in which considerations of data sufficiency require us to use monthly data. As we typically do not find strong evidence of dependencies among these results, the marginal effects of reporting date bias are likely negligible.

In the analysis undertaken in this study, we only used those loss data that were not listed as related to credit or market risk events, and where the net loss amount after direct recovery exceeded the reporting threshold of €20,000. In addition, the analysis used two versions of the database, according to the unit of measure being considered in the analysis. The full database was used in the analysis of losses by business line or by cell (combination business line / event type) category, but a "consolidated" database, in which loss records marked as "related" were combined into a single loss event,[1] was used in the analysis of correlations across event types.

[1]Losses marked as "related" in the database are losses stemming from a common root cause (event type) but having impacts in more than one business line of a bank.

| Business Lines | Event Types |
|---|---|
| 1. Corporate Finance | 1. Internal Fraud |
| 2. Trading and Sales | 2. External Fraud |
| 3. Retail Banking | 3. Employment Practices and Workplace Safety |
| 4. Commercial Banking | 4. Clients, Products, and Business Practices |
| 5. Clearing | 5. Disasters and Public Safety |
| 6. Agency Services | 6. Technology and Infrastructure Failures |
| 7. Asset Management | 7. Execution, Delivery, and Process Management |
| 8. Retail Brokerage | 8. Malicious Damage |
| 9. Private Banking | |
| 10. Corporate Items | |

Table 1: Primary business line and event type categories used in classifying ORX data.

In all, data on 90,024 loss events were used in the study, or 89,586 after related losses were consolidated. Because the number of related losses was so small with respect to the total number of losses, we did not consider the overall impact of related losses on the correlation values.

Not all banks reported losses in all categories, and in many cases, some banks only reported losses very infrequently in certain categories. In addition, different banks reported data over different time periods, with some providing a full six years' worth of data, while other members who recently joined the consortium only reporting a few months of data. To ensure that each time series considered had sufficient data to produce meaningful correlation and dependence measurements, we imposed the requirement that each time series include at least six periods worth of data (whether that period was annual, quarterly, or monthly), and that at least half of the time series values were nonzero. When comparing banks to each other, we only compared those loss categories in which there were sufficient loss numbers in each bank.

# 3 Scope and Background of the Analysis

In this section, we delimit the scope of the current analysis, and motivate our choices against the background of past studies into correlation, as well as against the availability of the consortium data. In reporting the results below, we were careful to maintain the strict levels of confidentiality regarding the identity of the consortium member banks. Therefore, we shall not present results on a bank-by-bank basis, but rather we report on the ranges of dependence measures observed across the consortium.

## 3.1 Granularity of Time Periods

Despite the rich amount of data on individual loss events from a variety of banks, the analysis was limited by the overall time period in which losses were reported. In order to perform the dependence analysis, the data for each bank were organized into aggregate loss time series computed at annual, quarterly, as well as monthly time scales, according to business line, event type, and cell categories. For example, for a given loss category, such as Retail Banking, the quarterly time series would record the total losses sustained in that business line for each quarter reported by the bank. Correlation and tail dependence measures were then applied to data in various categories by comparing data values occurring in the same period for the same bank.

Although the Basel regulations require dependence measures to be applied when aggregating estimates of loss distributions at an annual level from various units of measure, we unfortunately do not have sufficient data required to directly measure dependencies among annual losses for the banks in the consortium. Because the time series of annual data afforded at most six data points, few conclusive measures of annual loss dependence could be made. Monthly data, while more plentiful, often had issues of bias due to inconsistencies in reporting, with many losses reported only at the ends of quarters and ends of years. Quarterly data represents a compromise between reporting bias and data sufficiency. Therefore, when discussing measures of correlation, we shall generally report statistics based on quarterly data. However, we shall use monthly data for certain measures of tail dependence among the loss categories, as these require larger sample sizes to be meaningful. We shall discuss some of the issues involved with making inferences regarding annual correlations based on quarterly or monthly data later in this report.

### 3.2 Correlation and the LDA Framework

Correlation analysis is typically performed by banks as part of an overall application of the Loss Distribution Approach (LDA) method of computing capital charges. The most widely used framework to model operational risk, LDA is an actuarial model which quantifies OR on the base of identified internal loss events. Under this model, operational losses within various units of measure are supposed to be generated by two sources of randomness: their frequency and their severity (Chernobai, Rachev, and Fabozzi 2007). The total loss distribution of the bank over a fixed period of time (usually one year) is then determined in three stages:

1. Characterize the distribution of the frequency and severity of losses in each unit of measure,
2. Derive the distribution of the aggregate losses over the time period in each unit of measure,
3. Combine these aggregate loss distributions across the units of measure to arrive at an estimate of the total loss distribution of the bank, or at least determine certain quantiles of that distribution.

As mentioned in the introduction, one method of the computing of high quantiles of the total annual loss distribution in Step 3 of the LDA framework is performed by simply adding together the corresponding quantile estimates of the loss distributions for each unit of measure, which implicitly assumes a perfect correlation among the various units of measure. To avoid this assumption, banks usually introduce explicit dependence models at some point within the LDA framework. While there are many ways to introduce dependence, typically three methods are considered: (Frachot, Roncalli, and Salomon 2004; Boecker and Klueppelberg 2008)

1. The frequency distributions across multiple units of measure are dependent on each other;
2. The severities of individual losses in various units of measure are correlated, either through a direct modeling of dependence or through the introduction of "common shock" loss events affecting several units of measure simultaneously;

3. Dependence is introduced only at the level of the aggregate losses for each unit of measure.

Methods 1 and 3 are commonly used in practice, as they are the simplest to implement, and it is often difficult to formulate a coherent model that defines dependencies at the level of individual loss severities. Usually, measurements of frequency correlations are used to derive models for correlations at the aggregate loss level, while otherwise assuming independence of loss severities. Several authors, including (Nystrom and Skoglund 2002; Di Clemente and Romano 2003; Chapelle, Crama, Hubner, and Peters 2004; El Gamal, Inanoglu, and Stengel 2006), have proposed models in which correlations are only introduced at the aggregate loss level. However, other authors (Aue and Kalkbrenner 2006) point to some challenges for this approach if the effects of insurance mitigation are to be incorporated into this approach, which may add additional layers of dependence that are difficult to model through the straightforward application of a copula or correlation model. Both theoretical and empirical evidence indicate that correlations at the aggregate loss level are less, often much less, than the corresponding correlations among frequencies (Frachot, Moudaoulaud, and Roncalli 2003; Frachot, Roncalli, and Salomon 2004). As a result, in order to avoid confusion and to focus on the correlations of most direct relevance in computing capital charges, we shall only report on the dependence between aggregate losses in this report, where the aggregate losses incorporate the effects of direct loss recoveries, but omit the effects of indirect recoveries such as through insurance.

## 3.3 Copulas and Parametric Models of Dependence

Copulas have been widely discussed in the OR literature as a richer and more complete models of dependence between risk factors than what simple measures of correlation can provide. Copulas are multidimensional functions that are capable of fully expressing the nature of the dependence between two or more random variables; consequently they can be very complex mathematical objects. As a result, a variety of parametric forms for copulas have been studied, which have been designed to capture specific aspects of the joint distribution, such as tail dependence relations. We refer the reader to the discussions of copulas in (McNeil, Frey, and Embrechts 2005; Chernobai, Rachev, and Fabozzi 2007) for an introduction to the subject in the operational risk context. We just cite a few examples of the wide variety of applications of copulas to various aspects of dependence among operational losses: in (Chapelle, Crama, Hubner, and Peters 2004), the authors fit Gaussian and Frank copulas to frequencies; in (Dalla Valle, Fantazzini, and Giudici 2006), Gaussian and $t$-copulas are fitted to aggregate losses; in (Di Clemente and Romano 2003), the dependence structure among losses of different business lines is modeled using $t$-copulas; in (Wuthrich 2003), Archimedean copulas such as the Clayton, Weibull, and Gumbel copulas are employed as part of the estimation of the total capital charge; in (Embrechts and Puccetti 2006) Gaussian and Gumbel copulas are used. While copulas have become an important subject in dependence modeling, an evaluation of copula models with respect to various dependencies observed in the ORX data is out of the scope of the current study, and may form the basis of a future study. Instead, we shall try to capture tail dependence relations through direct, nonparametric empirical measures.

An alternative way to model dependence in operational loss data consists in using mixture models, an approach that has been commonly used in credit risk (see (Duffie and Singleton 1999) and (Lando 1998) for further references). Applications in the operational risk context are reported in (de Fontnouvelle and Rosengren 2004; Moscadelli 2004), among others. A recent application of mixture models for operational risk is reported in (Reshetar 2008). Here the author considers dependence among both frequencies and severities, resulting in dependence between the aggregate losses. Severities are modeled using an exponential-gamma mixture model, where losses from different classes of risk are driven by a common random parameter, and frequencies are similarly modeled using a Poisson-Gamma mixture model. In this study, parametric modeling of the frequency and severity distributions is beyond our scope, as again we shall be primarily concerned with nonparametric measures of correlation and dependence.

### 3.4 Serial correlation

Serial correlation refers to the correlation of losses within the same unit of measure across successive time periods. Serial correlation among operational losses has only been marginally addressed in prior studies. The operational risk literature, driven by the LDA model on one side (which automatically excludes serial correlations among severities) and copula theory on the other side (which is most natural in a static distributional context rather than a dynamic time series one (McNeil, Frey, and Embrechts 2005)), has focused on modeling dependencies more in a cross-sectional direction than in a longitudinal one. However, a few attempts to take it into account have been made in the literature. (Aue and Kalkbrenner 2006) observes that dependencies within a risk class should be generally higher than dependencies between different classes. Losses in the same business line/event type combination, for example, are often affected by similar internal processes and management structure, moreover occurring in the same business environment. We would expect a significant correlation among such losses. In some studies (see (Chernobai, Rachev, and Fabozzi 2007)), serial dependencies in frequency distributions are modeled through the use of Negative Binomial distributions or, more generally, non-homogeneous Poisson processes (i.e., Cox processes). Empirical evidence of seasonal and cyclic variation in loss frequencies have been also reported in (Chernobai, Burnecki, Rachev, Truck, and Weron 2006), (Allen and Bali 2005). Clearly, the identification of significant global trends and seasonalities in operational loss series can provide useful insights on common temporal patterns between risk classes, which should be taken into account in cross-sectional dependence considerations. At various points in the study findings below, we shall discuss the potential impact of serial correlation on our findings; however, we shall not provide a complete picture of serial dependence in the ORX data in this paper, but leave it as a topic for future exposition.

In the following three sections, we shall present our main findings: first, correlations among quarterly total losses at the business line, event type, and cell category level; second, measures of tail dependence among various units of measure, both pairwise and collectively; and third, direct empirical measures of the diversification benefits observed in the ORX data.

# 4 Correlation Analysis

We center our presentation of correlation results around quarterly aggregate loss data, using Kendall's $\tau$ as the measure of correlation. The choice of a rank correlation measure is appropriate for heavy-tailed aggregate loss data, as moment-based measures such as Pearson's correlation coefficient $\rho$ are highly unstable in such an environment. We prefer Kendall's $\tau$ over the alternative Spearman correlation coefficient as the latter measure is not associated with an intuitive measure of association in the population, and therefore point estimates and confidence intervals for Spearman's $\rho$ are not particularly meaningful (Hollander and Wolfe 1999). We applied the standard estimator of $\tau$, which counts the number of instances of data pairs in which two data sets agree in the order of their value, minus the number of instances of pairs where the data sets disagree in their ordering: ((Hollander and Wolfe 1999))

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}[(X_i - X_j)(Y_i - Y_j)],$$

where

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Note that ties in the data are not counted in the terms in the sum; one result of this is that the correlation measure applied between a time series and itself can be less than one in value if ties are present in the data.

Prior studies in the banking context have indicated that low values should be expected when measuring correlations of aggregate loss values. For example, (Frachot, Moudaoulaud, and Roncalli 2003; Frachot, Roncalli, and Salomon 2004) present analytical arguments to indicate that the range of Pearson (linear) correlation values should be no higher than 0.1. In a study of

3,000 losses collected from a large European bank, (Chapelle, Crama, Hubner, and Peters 2004) find Spearman correlation values among loss frequencies that are typically less than 0.3, but ranging up to near 0.6 in some cases; as (Frachot, Moudaoulaud, and Roncalli 2003) explain, the corresponding correlation values among aggregate loss values will be less than those of the frequencies. (El Gamal, Inanoglu, and Stengel 2006) find low values (typically less than 0.1) of the Pearson correlation coefficient for weekly loss data taken from one of the participating banks in the 2004 Loss Data Collection Exercise (cf. (Federal Reserve System 2005)), where some outlying data points were removed that were causing instability in the linear correlation measure.

We also observe low correlation values in the ORX data. Figure 1 shows the range of values observed for selected business line pair among all banks reporting losses in both those categories.[2] The numbers of banks reporting losses in each business line category pair is presented in Table 2. We see that in almost all cases (with the exception of Retail Brokerage / Corporate Items), the median correlation value for each business line pair is less than 0.2. Overall, we rarely see correlation values above 0.4. In addition we see quite a few negative correlation values; this can be expected due to sampling error if the true correlation is near zero. The mean value of all the correlation values among all business line categories was 0.062, with a standard deviation of 0.178. We note that the theoretical standard deviation of the sample if all the loss categories were independent of each other would be 0.15 for a bank reporting 24 quarters of losses and 0.18 for a bank reporting 16 quarters of losses; of the 34 banks in the sample, 29 reported at least 16 quarters of losses.

Figure 2 shows a similar range of values for event type pairs. In this figure, we see a slightly smaller range of observed values as with business line data, and we observe that only in the case of Disasters and Public Safety/Malicious Damage that the median correlation value across banks exceeds 0.2, although it should be mentioned that there are only three banks reporting sufficient losses in this category; see Table 3, which reports the numbers of banks reporting data for each selected category pair. The mean value of all the correlation values among all event type categories was 0.059, with a standard deviation of 0.188.

In addition, Figure 3 shows a histogram of all average correlation values observed among losses in all pairs of cell categories (Since there are 3,160 different pairs that may be formed from the 80 cells (i.e., combinations of business lines and event types), we could not practically show the range of values observed for each cell category pair.) We see from this histogram that the observed correlation values appear to be very nearly normally distributed, with a mean value of 0.056, and a standard deviation of 0.171.

[2] In this paper, we only present detailed correlation values for selected business lines and event types. Full results are available to ORX consortium member banks.

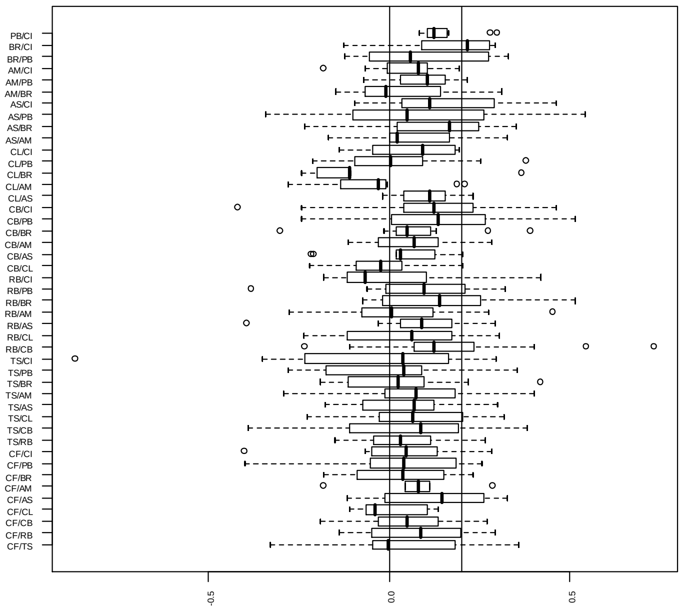


Figure 1: Ranges of selected values observed among within-bank correlations of quarterly aggregate loss values by business lines. The boxplots are drawn in the standard manner, and are composed from the point correlation estimates of each bank reporting losses in the business line pair. Vertical lines are drawn at the 0 and 0.2 correlation levels to indicate a general *a priori* range in which correlations were expected to fall. The abbreviations CF, TS, RB, CB, CL, AS, AM, BR, PB, and CI stand for Corporate Finance, Trading and Sales, Retail Banking, Commercial Banking, Clearing, Agency Services, Asset Management, Retail Brokerage, Private Banking, and Corporate Items, respectively.

| | CF | CB | CL | AS | AM | BR | PB | CI |
|---|---|---|---|---|---|---|---|---|
| CF | 8 | 8 | 5 | 4 | 6 | 4 | 8 | 7 |
| CB | | 28 | 12 | 9 | 17 | 11 | 18 | 14 |
| CL | | | 14 | 7 | 11 | 5 | 13 | 7 |
| AS | | | | 10 | 8 | 7 | 9 | 6 |
| AM | | | | | 18 | 8 | 15 | 9 |
| BR | | | | | | 12 | 10 | 6 |
| PB | | | | | | | 20 | 11 |
| CI | | | | | | | | 14 |

Table 2: Numbers of banks reporting sufficient losses in each business line pair. Numbers along the diagonal represent the number of banks reporting sufficient losses in that business line.

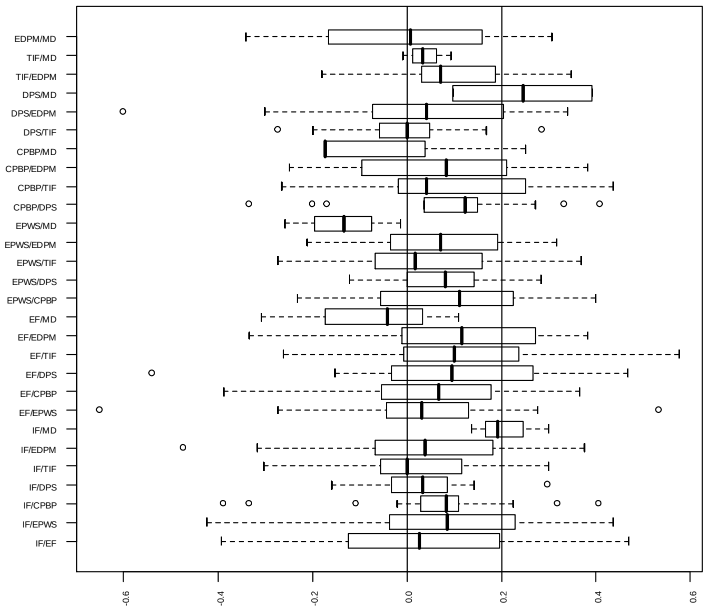


Figure 2: Ranges of selected values observed among within-bank correlations of quarterly aggregate loss values by event types. The boxplots are drawn in the standard manner, and are composed from the point correlation estimates of each bank reporting losses in the business line pair. Vertical lines are drawn at the 0 and 0.2 correlation levels to indicate a general *a priori* range in which correlations were expected to fall. The abbreviations IF, EF, EPWS, CPBP, DPS, TIF, and EDPM stand for Internal Fraud, External Fraud, Employment Practices and Workplace Safety, Clients Products and Business Practices, Disasters and Public Safety, Technology and Infrastructure Failures, and Execution Delivery and Process Management, respectively.

| | IF | EPWS | DPS | TIF | EDPM | MD |
|---|---|---|---|---|---|---|
| IF | 26 | 22 | 12 | 18 | 26 | 3 |
| EPWS | | 24 | 13 | 17 | 24 | 3 |
| DPS | | | 13 | 9 | 13 | 2 |
| TIF | | | | 19 | 19 | 3 |
| EDPM | | | | | 34 | 3 |
| MD | | | | | | 3 |

Table 3: Numbers of banks reporting sufficient losses in each event type pair. Numbers along the diagonal represent the number of banks reporting sufficient losses for that event type.

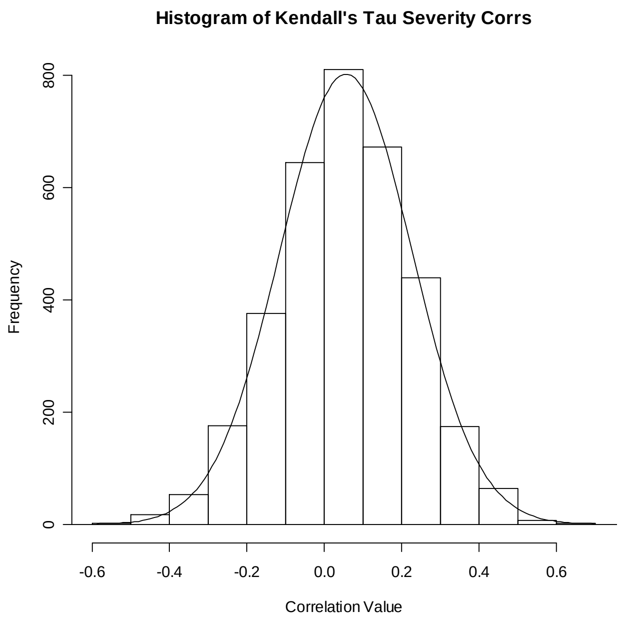


Figure 3: Histogram indicating the range of within-bank correlation values observed among quarterly aggregate loss values by cell combinations. The line shows the density function of a normal distribution with mean and standard deviation equal to the sample mean and sample standard deviation, which are 0.056 and 0.171, respectively.

The boxplots and histograms indicate that most correlation values are similar. In the following tables, we report the average correlation values observed for each business line pair and event type pair. In particular, we note that higher correlation values are observed in the Corporate Items and Malicious Damage category, although the low sample size of the latter category should again be noted.

We applied a chi-square test of equality between each bank's correlation matrix and the corresponding submatrix of the average correlation matrices reported in Tables 4 and 5. The test proceeds as follows: we first construct a matrix $\Gamma$ of the estimated variances and covariances among all correlation values observed between each pair of business lines or event types. In the case of business lines, where there are up to 10 categories and 45 estimated correlation values for each bank, $\Gamma$ would be a 45-by-45 variance-covariance matrix. The variances and covariances among the Kendall correlation values were computed following the formulas of (Cliff and Charlin 1991). A chi-square test of homogeneity between the correlation values was applied by computing the statistic

$$(x - \overline{x})^T \Gamma^{-1} (x - \overline{x}), \tag{1}$$

where $x$ is a vector of estimated Kendall correlation values, $\overline{x}$ is the corresponding vector of average correlation values, and $\Gamma^{-1}$ is the inverse of the estimated variance-covariance matrix. Under the null hypothesis that the true correlation values are given by $\overline{x}$, (1) has an approximate $\chi^2$ distribution with $p(p-1)/2$ degrees of freedom, where $p$ is the number of columns in the bank's estimated correlation matrix. In several cases, however, with the amount of data available quarterly data, the sample-based estimates of the matrix $\Gamma^{-1}$ were unstable, often leading to matrices that were not positive definite, and yielding negative test statistics. Therefore, we ignored the covariance terms in the matrix and simply treated $\Gamma$ as a diagonal matrix whose positive elements were estimates of the variance of each correlation measure, using the method of (Cliff and Charlin 1991).

The results of applying this test at a 0.05 significance level are displayed in Table 6. We applied a Bonferroni adjustment to the $p$-values of the individual tests to accommodate the multiple comparisons across the test levels. We see that in a majority of cases (84% for business lines and 88% for event types), each bank's sample correlation matrix appeared to be within the range of sampling error of the values given in the average correlation tables. This indicates that the average correlation matrix may be used with confidence to represent the correlations observed among ORX consortium members. However, we should note that the test lacks some power due to the low number of quarterly data samples available to us from most banks.

| | CB | CL | AS | AM | BR | PB | CI |
|---|---|---|---|---|---|---|---|
| CF | 0.05 | 0.01 | **0.13** | 0.07 | 0.03 | 0.03 | 0.01 |
| CB | | -0.02 | 0.03 | 0.06 | 0.07 | **0.13** | **0.11** |
| CL | | | **0.10** | -0.04 | -0.06 | 0.02 | 0.06 |
| AS | | | | 0.07 | **0.12** | 0.06 | **0.15** |
| AM | | | | | 0.04 | 0.09 | 0.05 |
| BR | | | | | | **0.10** | **0.16** |
| PB | | | | | | | **0.15** |

Table 4: Average Kendall correlation values observed for quarterly aggregate loss data among business lines. Values exceeding 0.1 are highlighted in bold.

| | EPWS | DPS | TIF | EDPM | MD |
|---|---|---|---|---|---|
| IF | 0.07 | 0.04 | 0.00 | 0.04 | **0.21** |
| EPWS | | 0.07 | 0.04 | 0.07 | -0.14 |
| DPS | | | 0.00 | 0.01 | **0.24** |
| TIF | | | | **0.10** | 0.04 |
| EDPM | | | | | -0.01 |

Table 5: Average Kendall correlation values observed for quarterly aggregate loss data among event types. Values exceeding 0.1 are highlighted in bold.

| | Business Lines | Event Types |
|---|---|---|
| Number of banks tested | 32 | 33 |
| Number where test passes | 27 | 29 |
| Percent where test passes | 84% | 88% |

Table 6: Outcomes of the chi-square tests of homogeneity of each bank's quarterly aggregate loss correlation matrix against the average values displayed in Tables 4 and 5.

## 4.1 Annual vs. Quarterly Correlation Measures

Due to the lack of annual data, we have focused instead on presenting the results on correlations among quarterly aggregated time series. Unfortunately, there are no simple relationships from which one may infer annual correlations from correlations of quarterly or monthly data. This can be illustrated analytically using the Pearson correlation coefficient, as similar phenomena have been observed using Kendall's $\tau$ in simulated data. For simplicity of exposition, let us suppose we have two values of one time series $X_1$, $X_2$ and two values of another time series $Y_1$, $Y_2$, each representing half-year aggregate loss data; the corresponding annual data from each series is $X_1 + X_2$ and $Y_1 + Y_2$. Assuming that both time series are stationary, the Pearson correlation between $X$ and $Y$ is given by

$$\mathrm{Cor}(X_1 + X_2, Y_1 + Y_2) = \frac{2\mathrm{Cov}(X,Y) + \mathrm{Cov}(X_1, Y_2) + \mathrm{Cov}(X_2, Y_1)}{2\sqrt{\sigma_X^2 + \mathrm{Cov}(X_1, X_2)} \cdot \sqrt{\sigma_Y^2 + \mathrm{Cov}(Y_1, Y_2)}}.$$

When there is no serial correlation in either time series, and there is no cross-correlation at lagged values between the series $X$ and $Y$, then the last two terms in the numerator and the covariance terms in the denominator are all zero, and the annual correlations equal the half-year correlations between $X$ and $Y$. If, however, there is positive autocorrelation in one time series but not in the other, and there is otherwise no cross-lagged correlation, then the value of the denominator increases and the annual correlation is smaller than the quarterly correlation. This case might occur in practice if, for example, one time series exhibits an increasing trend behavior and the other does not. If however both time series have a positive serial correlation, such as may be observed when both series have a similar trend, then the annual correlations may well be higher than the half-year correlations, and we have observed in both simulated and actual data that this often occurs.

We measured the serial correlation effect using the autocorrelation values of quarterly ranked losses among all banks in the sample, for both business lines and event types. In each case, we found similar results. Among 182 reported quarterly business line time series, the lag-1 autocorrelation (measured using Kendall's correlation coefficient) was 0.044, with a standard deviation of 0.176. Among 179 reported quarterly event type time series, the mean was 0.048 and standard deviation was 0.189. In each case, the distribution of sample autocorrelations was closely matched by a Normal distribution. For both business line and event type autocorrelations, we tested the hypothesis that the autocorrelations for each bank were equal to the mean autocorrelation value, and found that both tests passed at the 0.05 significance level. A similar test of independence found that only one measured autocorrelation value was significantly different from zero (at the 0.05 significance level). Therefore we did not find any strong evidence of pervasive or large serial correlations among the quarterly business line and event type time series.

# 5 Tail Probability Measures and Simultaneous Exceedances

## 5.1 Coefficient of tail dependence

Due to the heavy-tailed nature of the data, extreme losses tend to be several orders of magnitude greater than median losses, and can be very severe. When a business line incurs an extreme loss, this loss will typically overshadow all other losses from other business lines, unless of course another business line also experiences an extreme loss. When multiple extreme events occur, the results can be especially devastating. Such events of course can occur by sheer chance, but the presence of *tail dependencies* in the loss distributions across more than one loss category can result in such events occurring more frequently than would otherwise be expected. In this section, we shall investigate the likelihood with which extreme losses occur simultaneously in different loss categories, and determine if this likelihood is greater or less than what one might expect if losses were realized independently across business lines.

The coefficient of tail dependence $\lambda$ is a measure of the dependencies between extreme values of a bivariate distribution. That is, for two random variables $X$ and $Y$, the coefficient provides an indication of how likely it is to realize an extreme value of $Y$ given that $X$ takes an extreme value. Formally, it is defined as a limit of the distribution function:

$$\lambda(X, Y) = \lim_{q \to 1} P(Y > Q_Y(q) \mid X > Q_X(q)),$$

where $Q_X(q)$ is the quantile function of the distribution of $X$ evaluated at the probability level $q$. If the limit exists and is positive, then there is said to be a tail dependence between $X$ and $Y$. As was true of the rank-based correlation measures, the coefficients of tail dependence depend only on the copula of a bivariate distribution. Additional details can be found in (McNeil, Frey, and Embrechts 2005).

We shall focus on an empirical measure of tail dependence, which we shall term the *tail ratio*. This ratio is defined for series $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, for a given $q \in (0, 1)$, as

$$\begin{aligned} \hat{\lambda} &= \frac{\#_i\{X_i > \hat{Q}_X(q), Y_i > \hat{Q}_Y(q)\}}{\#_i\{X_i > \hat{Q}_X(q)\}(1-q)}, \\ &= \frac{\#_i\{X_i > \hat{Q}_X(q), Y_i > \hat{Q}_Y(q)\}}{\lfloor n(1-q) \rfloor \cdot (1-q)}, \end{aligned}$$

where $\hat{Q}_X(q)$ represents the quantile function of the empirical distribution function for $X_1, \ldots, X_n$, and $\#\{\cdot\}$ represents the count of the data values satisfying the condition in brackets. Note that if the tails of the distributions of $X$ and $Y$ are independent, then the expected value of $\hat{\lambda}$ equals one. If the ratio exceeds one, this may indicate a possible positive tail dependence between the two distributions, such that extreme values of $Y$ are more likely to be realized if $X$ takes an extreme value. The value of the tail ratio may be considered to be the factor multiple by which the observed likelihood exceeds the expected likelihood, under the assumption of independence, that losses from both categories simultaneously exceed the quantile level $q$. For example, let us say that losses in a category are "large" if they exceed a certain quantile level $q$ of their distribution. If the tail ratio value between loss categories A and B is 2, that indicates that losses observed in category A are twice as likely to be large when losses in category B are large than they would be if the two loss categories were tail-independent.

The distribution of $\hat{\lambda}$ is easy to determine under the assumption of tail independence, as the number of $Y$ data values exceeding the $q$th quantile of its distribution when $X$ exceeds its $q$th quantile is binomially distributed with parameters $\lfloor n(1-q) \rfloor$ and $(1-q)$. Based on this observation, we can test for tail independence between two loss distributions by checking if the tail ratio falls above or below a given interval around one.

We have chosen the value of the quantile parameter $q$ to equal 0.8 throughout this study, as this choice balances the need to focus on the tails of the distributions with the limited amount of data available to estimate the tails. We shall report findings based on monthly data in this section, as there can be up to 72 data points available, which implies an expected number of simultaneous exceedances of the 0.8 quantile of $\lfloor 72(0.2) \rfloor \cdot 0.2 = 2.8$. With this amount of data, a test of tail independence would reject the null hypothesis at a significance level of approximately 0.13 if the tail ratio exceeds 1.4, or at a significance level of 0.04 of the tail ratio exceeds 1.8. (Not all banks report 72 monthly loss values, and so the interval between these threshold values generally undercovers the true confidence region, implying that the tests tend to fail more easily.) The following charts indicate the range of tail ratios observed on monthly aggregate loss data across all banks for each pair of business lines (Figure 4) or event types (Figure 5). Only in Figure 4, in business line data, do we see one category (Corporate Finance / Clearing) where the median tail ratio exceeds the 0.04 significance threshold. Of course, given the number of categories being compared, these excessive tail ratios may arise by chance. Among event types, however, all the median tail ratios are well within the confidence limits.

Figure 6 indicates the histogram of all tail ratios observed at the cell level across all banks. The solid line indicates the values that would be expected (approximately) if the loss categories were independent and followed the binomial distribution described earlier. Across the range of tail ratio values, we see that there is a slightly higher frequency of large tail ratios than what one might expect under the independence hypothesis.

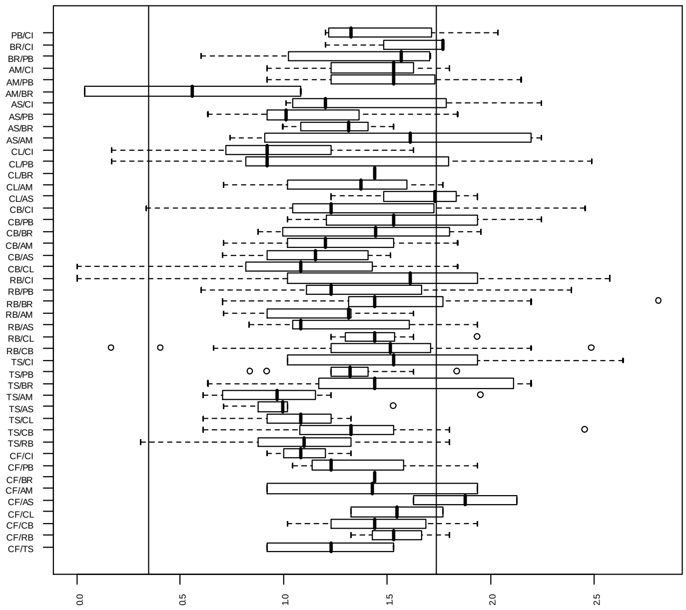


Figure 4: Ranges of values observed among tail ratios of monthly aggregate loss values by business lines. The vertical lines represent approximate critical values of a two-sided test of tail independence at at 0.08 significance level.

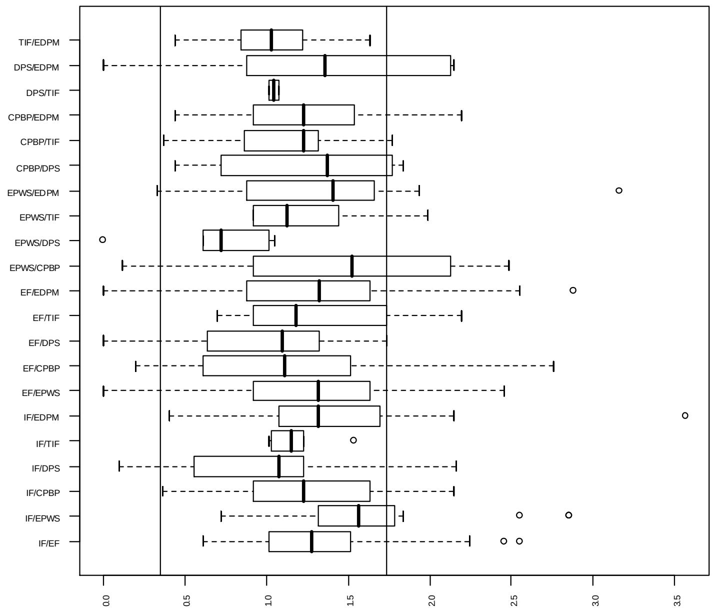


Figure 5: Ranges of values observed among within-bank correlations of quarterly aggregate loss values by event types.

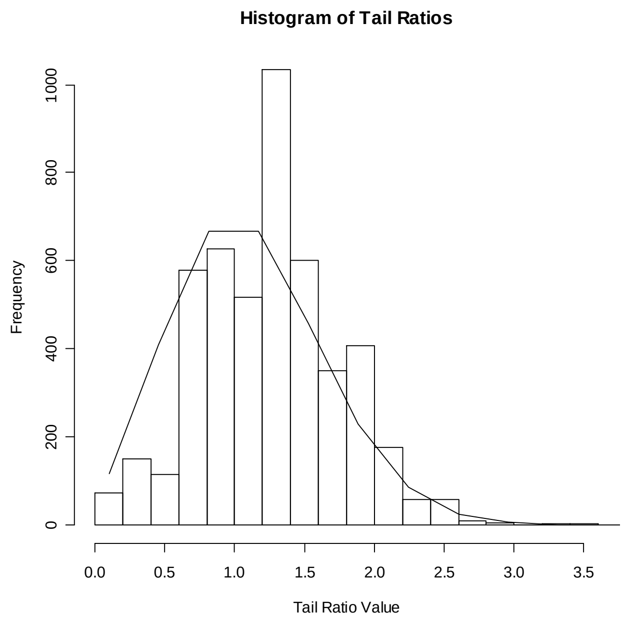


Figure 6: Histogram indicating the range of tail ratios observed among monthly aggregate loss values by cell combinations. The line indicates the probability mass function of a binomial distribution that approximately corresponds to the expected distribution of the data under the independence hypothesis.

Next, we turn away from pairwise measures of tail dependence between two loss categories, and consider the number of simultaneous exceedances of the 0.8 quantile level reported by banks across all business lines, event types, or cells occurring during the periods in which they have reported data to ORX. These measures were computed as follows. For each bank, within each loss category, the 0.8 quantile level of the total losses reported in all periods was computed. Then, for each period in which the bank reported losses, we computed the number of categories in which the total losses exceeded the corresponding 0.8 quantile estimate in that category. For example, suppose a bank only reports losses in Retail Banking, Commercial Banking, and Trading and Sales, and suppose that the 80th percentile losses reported by that bank in each of these business lines across all quarters was €800k, €1.1m, and €1.5m, respectively. Suppose in a given quarter, the bank reported losses of €1m, €700k, and €2m. Here, both the Retail Banking and Trading and Sales losses exceeded the estimated 0.8 quantile level, and we would record 2 simultaneous exceedances.

Under the hypothesis that there is no tail dependence across loss distributions in each category, the distribution of the number of simultaneous exceedances is binomial, with parameters $(n, p)$ given by the number of loss categories and the exceedance probability 0.2, respectively. In the following figures, we present bar charts that display the counts of simultaneous exceedance values observed across all banks for business lines (Figure 7), event types (Figure 8), and cells (Figure 9), based on quarterly data. Beside the bars indicating the observed counts of exceedance values, we also plot bars indicating the expected total number of counts for those numbers of simultaneous exceedances, which were obtained by simply summing up the probabilities of achieving those counts over each reporting period for each bank.

These charts indicate that at quarterly timescales, the number of simultaneous exceedances observed is typically very close to the expected number if the business lines are tail-independent, although in each case we see that there are some instances where larger numbers of simultaneous exceedances have been observed to occur, over what might be expected under tail-independence. In the case of business lines, there was one observation each of 6 and 7 simultaneous exceedances among banks; the expected values are very close to zero. In the case of cell categories, on four occasions we observed 13 simultaneous exceedances, and we observed each of 14 and 15 simultaneous exceedances twice in the sample, and one time we observed 20 simultaneous exceedances. These do appear to be significant, although rare outliers in the sample.

## 6 Diversification Benefits

To this point, we have seen that correlations and tail dependencies among losses in various business lines, event types, and cell categories are not particularly strong. We now ask whether

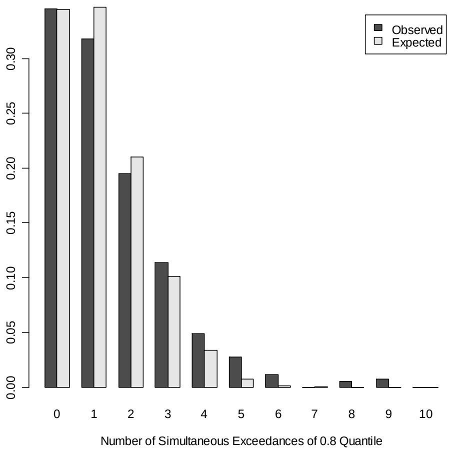


Figure 7: Counts of the numbers of simultaneous exceedances of total business line losses of the 0.8 quantile level, observed across all banks across all quarters.

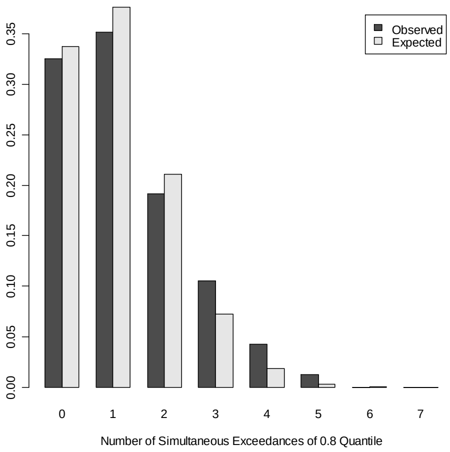


Figure 8: Counts of the numbers of simultaneous exceedances of total business line losses of the 0.8 quantile level, observed across all banks across all quarters.

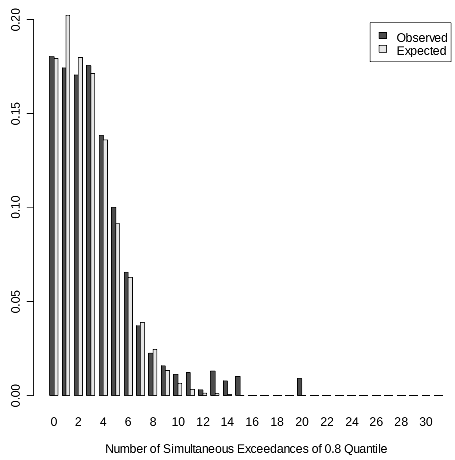


Figure 9: Counts of the numbers of simultaneous exceedances of total cell losses of the 0.8 quantile level, observed across all banks across all quarters.

these findings have any bearing on the computation of capital charges using the Loss Distribution Approach (LDA), which is perhaps the most widely used method by banks in fulfillment of the AMA requirements of Basel II. As it is currently mandated by many regulators, the default accepted method of handling dependence within the LDA framework, in the absence of suitable correlation data, is to estimate a high quantile of the total loss distribution as the sum of estimates of the corresponding quantiles of the annual loss distribution of each of several units of measure. In other words, the 99.9th percentile of the sum of all losses across all categories is computed as the sum of the 99.9th percentiles of each loss category. Usually, this approach to computing capital requirements is viewed as being highly conservative, since this computation would be correct if all loss categories exhibit perfect correlation with each other, so that high losses in one loss category will co-occur with high losses in every other loss category.

Our correlation results indicate that we are far from the situation of perfect correlation, and this may indicate that there are benefits to be had from the diversification of the bank's operational losses across the various loss categories. A diversification benefit would imply that high quantiles of the total annual loss distribution would be much less than the sum of the corresponding quantiles of the annual loss distributions from each category. Many authors have provided empirical evidence about the diversification effect and its impact on capital charge calculations (see, for example, (Aue and Kalkbrenner 2006; Chavez-Demoulin, Embrechts, and Neslehova 2005; Dalla Valle, Fantazzini, and Giudici 2006; Di Clemente and Romano 2003; El Gamal, Inanoglu, and Stengel 2006; Frachot, Moudaoulaud, and Roncalli 2003; Frachot, Roncalli, and Salomon 2004; Powojowski, Reynolds, and Tuenter 2002; Reshetar 2004; Reshetar 2008)). However, theoretical studies have shown that in the presence of heavy-tailed loss distributions, there may be no diversification effect or worse, some quantiles of the total loss distribution are even greater than the sum of the loss category quantiles (Neslehova, Embrechts, and Chavez-Demoulin 2006).

We can demonstrate this effect through a simple example. Suppose a bank has two business lines, such that each business line loses €100k $100(1-\alpha)\%$ of the time, and €1m $100\alpha\%$ of the time. In this case, the sum of $100(1-\alpha)\%$ percentile losses for each business line is €200k. The total losses in any period may be €200k, €1.1m, or €2m, depending on whether zero, one, or both business lines experience high losses. In the case where the business lines are perfectly correlated (i.e., Kendall's $\tau = 1$), the $100(1-\alpha)\%$ percentile of total losses is 200k, as implied by the perfect correlation approach. However, for any value of Kendall's $\tau < 1$, the $100(1-\alpha)\%$ percentile of total losses is 1.1k, which is 5.5 times greater than the sum of the corresponding quantiles of the individual business lines. This simple example shows that with heavy tails, it is not unexpected to have a reverse diversification effect. The chance that at least one business line has a greater than €1m loss exceeds the chance that any given business has a greater than

€1m loss. Also, if any business line has a greater than €1m loss, then that loss dominates the sum of all other losses by an order of magnitude. Therefore, the quantiles of the total losses will dominate the sum of the corresponding quantile of the losses in each business line. This effect is independent of the degree of correlation among business lines.

To determine if diversification effects were apparent in the aggregate loss values reported by the banks in the ORX database, we empirically measured quantiles of the total loss distribution (the empirical quantile of the sum of losses in each business line category) and compared it to the sum of the corresponding VaR of each business line category for each bank. The ratio of these two values was plotted for each bank at every quantile level between 0.85 and 1. Because we based the quantile estimates on the empirical distribution functions, the quantile level 1 returns the maximum observed value in the dataset. A plot of these ratios for quarterly data can be found in Figure 10.

Figure 10 indicates that at the higher end of the quantile range, we see some evidence of diversification benefit. At all the displayed quantile levels, most banks show a ratio below one, and at the 0.96 level, nearly all the banks show evidence of diversification benefit. Although we are likely safe in assuming that a diversification benefit will continue to be realized at quantile levels above 0.96, there is no guarantee of this, nor can we reliably estimate what the diversification benefit would be at the 0.999 quantile level, as specified by the regulatory agencies. On the other hand, at quantile levels below 0.9, the ratios exceed one for many banks, indicating superadditivity of the quantile values at these quantile levels. We emphasize however that this effect does not indicate the presence of "hidden dependencies" among business line data that have not been observed in the above studies of correlation and tail dependence. Rather, it is primarily an artifact of the heavy-tailed nature of the loss distributions and the non-coherence of the quantile level (VaR) as a measure of risk. This indicates that the tail behavior of the loss distributions can play (at least) as important a role as the degree of correlation or dependence in determining the level of diversification benefit across business lines.

# 7 Factors Influencing Correlation and Dependence

We conclude this report by mentioning some of the limitations to which this analysis was subject due to the nature of the ORX data. First, as we have discussed earlier, the sample sizes for computing correlations at an annual level are very small, as there are at most six years of data available from any member bank. Annual data is of greatest interest to members as capital charges must be calculated to cover the losses over a one year period with high probability. The lack of data has led us in many cases to turn to quarterly and monthly data instead as a proxy for annual data. As the discussion in Section 4.1 indicated, there is no simple formula that allows

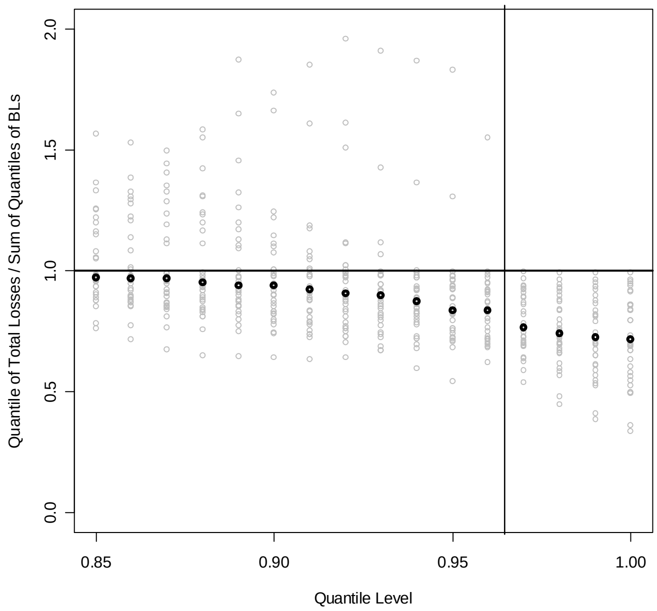


Figure 10: Ratios of the quantiles of the total losses divided by the sum of the corresponding quantiles of each business line, for quantile levels between 0.85 and 1. The graph shows the ratios based on quarterly data. Each gray point represents the value of the ratio for one bank. The larger black points indicate the median value across all banks. The points at the far right at quantile level 1 represent the maximum total loss divided by the sum of the maximum loss in each business line, which is mathematically guaranteed to be less than one. Because at most 24 quarterly loss values are reported by each bank, the ratio values at quantile levels between 0.96 and 1 represent an interpolation of those values. The points to the right of the vertical line therefore should not be considered as strong evidence of the diversification benefit above the 0.96 level. However, we include these points in the graph to indicate a lower bound estimate on the range of reductions that may be possible with diversification at high quantiles.

one to derive annual correlations from quarterly correlations in such a way that they would be more accurate than a direct measurement of annual correlations would be. Also, whether annual correlations are larger or smaller than quarterly correlations has much to do with the presence of trends or other types of serial correlation in the data.

Trends can be observed in many of the time series of loss values used in the correlation analysis in this study. Trends among a single bank's loss categories may be attributable to a number of factors, including bank growth, mergers and acquisitions, improvements or changes in the loss reporting and the control environment, and the completeness of reporting among recently incurred losses. Other factors that may affect several banks simultaneously include changes in the legal and regulatory environment, and, importantly, fluctuations in the exchange rate, which we discuss in more detail below. Trends are associated with longer-term fluctuations in the frequency, severity, or aggregate loss values experienced by a loss category, and often can be predicted with high accuracy, unlike the random "shocks" that can unexpectedly create large losses in one or more units of measure. One approach to dealing with trends would be to measure correlations both on series containing trends, as well as on the same series where an estimated trend has been removed. The correlations among the residuals can provide information on the degree to which the shocks experienced by the banks occur simultaneously to more than one unit of measure. In addition, the estimated trends themselves can be compared with trends in other units of measure within the same bank, or the same unit of measure within different banks, to determine the extent to which long-term changes in loss values are influenced by bank-specific or by industry-wide effects.

As mentioned above, a potentially significant source of correlation arises because much of the ORX loss data have been subject to fluctuating exchange rates that affect both the frequency and severity levels of losses incurred by banks in currencies other than the Euro. In particular, the US dollar lost 20% of its value to the Euro during the two-year period from 2006 to 2007. This implies that losses of equal value incurred by a bank in the United States at the beginning of 2006 and at the end of 2007 would appear to have a 20% difference in value in the ORX database. Furthermore, as a result the ORX reporting threshold of €20k, the decline in the dollar's value implies that a loss whose value just exceeded the reporting threshold in 2006 might not have exceeded the threshold in 2007. Given the large number of loss values observed in the €20-25k range, the frequency at which losses were reported to occur in 2007 is consequently much less. (However, this would affect correlations among frequency values much more than correlations among aggregate loss values.) The decline in frequency and severity of losses implies that the aggregate losses from US banks tend to follow a decreasing trend over recent years, and this trend will affect many business lines simultaneously.

## 8 Concluding Summary

In summary, this study has shown that:

- Most of the correlations among quarterly aggregate losses are low, generally less than 0.2, and rarely exceeding 0.4. Moreover, the correlation structures of individual banks appear to be largely homogeneous. A formal statistical test for the equality of correlation matrices indicated that the majority of individual banks' correlation matrices were found to be statistically equal to the average correlation matrix. Therefore, the average correlation matrix is representative of the correlation of most ORX members.

- Measurements of tail dependence show little evidence of tail dependence among units of measure. Although we observe a slightly elevated occurrence of simultaneous exceedances over the 80th percentile among various units of measure, most of these measurements were in line with what would be expected under the assumption of tail independence.

- Direct empirical measures of diversification benefits found evidence that such benefits can be realized at the higher end of the quantile range, although we do not have enough data to provide accurate estimates of this benefit at the 0.999 quantile level. However, despite the observed lack of strong correlation and tail dependence, at lower quantile values a reverse diversification effect was found, which is an artifact of the heavy-tailed nature of the loss distributions and the non-coherence of the quantile level (VaR) as a measure of risk. This indicates that the tail behavior of the loss distributions plays as important a role as the degree of correlation or dependence in determining capital requirements.

## References


Allen, L. and T. G. Bali (2005). Cyclicality in catastrophe and operational risk measurements. Department of Economics and Finance, Baruch College, New York, NY. Available at `http://www.gloriamundi.org/picsresources/latb_0511.pdf`.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

Aue, F. and M. Kalkbrenner (2006). LDA at work. Risk Analytics and Instruments, Risk and Capital Management, Deutsche Bank, Frankfurt. Available at `http://www.gloriamundi.org/picsresources/famk_lda_v2.pdf`.

Basel Committee for Banking Supervision (2003). The New Basel Capital Accord – Third Consultative Paper.

Boecker, K. and C. Klueppelberg (2008). Modeling and measuring multivariate operational risk with Lévy copulas. *Journal of Operational Risk 3*, 3–27.

Chapelle, A., Y. Crama, G. Hubner, and J. P. Peters (2004). Basel II and operational risk: implications for risk measurement and management in the financial sector. NBB Working Paper No. 51 May 2004, National Bank of Belgium. Available at `http://www.gloriamundi.org/picsresources/cchp.pdf`.

Chavez-Demoulin, V., P. Embrechts, and J. Neslehova (2005). Quantitative models for operational risk: extremes, dependence and aggregation. Presented at Federal Reserve Bank of Boston May 18-20 2005. Available at `http://www.gloriamundi.org/picsresources/vcdpejn.pdf`.

Chernobai, A., K. Burnecki, S. T. Rachev, S. Truck, and R. Weron (2006). Modeling catastrophe claims with left-truncated severity distributions. *Computational Statistics 21*(3), 537–555.

Chernobai, A. S., S. T. Rachev, and F. J. Fabozzi (2007). *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis.* Wiley Finance.

Cliff, N. and V. Charlin (1991). Variances and covariances of Kendall's Tau and their estimation. *Multivariate Behavioral Research 26*, 693–707.

Dalla Valle, L., D. Fantazzini, and P. Giudici (2006). Copulae and operational risks. Department of Statistics, University of Milano-Bicocca, Italy. Available at `http://www.gloriamundi.org/picsresources/lddfpg_cao.pdf`.

de Fontnouvelle, P. and E. Rosengren (2004). Implications of alternative operational risk modeling techniques. Working paper, Federal Reserve Bank of Boston. Available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=556823`.

Di Clemente, A. and C. Romano (2003). A copula-extreme value theory approach for modeling operational risk. Department of Economic Theory and Quantitative Methods for the Political Choices, University of Rome.

Duffie, D. and K. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies 12*, 687–720.

Dugan, J. C., J. J. Johnson, R. E. Feldman, and J. M. Reich (2007). Proposed supervisory guidance for Internal Ratings-Based systems for credit risk, Advanced Measurement Approaches for operational risk, and the supervisory review process (Pillar 2) related to Basel II implementation. Department of the Treasury, Office of the Comptroller of the Currency, Docket No. OCC-2007-0004. Available at `http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20070215a1.pdf`.

El Gamal, M., H. Inanoglu, and M. Stengel (2006). Multivariate estimation of operational risk with judicious use of extreme value theory. Economics Working Paper WP2006-3, Office of the Comptroller of the Currency. Available at `http://www.gloriamundi.org/picsresources/meghims_meo.pdf`.

Embrechts, P. and G. Puccetti (2006). Aggregating risk capital, with application to operational risk. Department of Mathematics, ETH Zurich. Available at `http://www.gloriamundi.org/picsresources/pegp_1.pdf`.

Federal Reserve System (2005). Results of the 2004 Loss Data Collection Exercise for Operational Risk. Available at `http://www.bos.frb.org/bankinfo/qau/pd051205.pdf`.

Frachot, A., O. Moudaoulaud, and T. Roncalli (2003). Loss distribution approach in practice. Working paper, Groupe de Recherche Operationnelle, Credit Lyonnais, France. Available at `http://www.gloriamundi.org/picsresources/afomtr.pdf`.

Frachot, A., T. Roncalli, and E. Salomon (2004). The correlation problem in operational risk. Working paper, Groupe de Recherche Operationnelle, Credit Lyonnais, France. Available at `http://www.gloriamundi.org/picsresources/aftres.pdf`.

Hollander, M. and D. A. Wolfe (1999). *Nonparametric Statistical Methods*. Wiley. 2nd ed.

Lando, D. (1998). Cox processes and credit risk securities. *Review of Derivatives Research 2*, 99–120.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management*. Princeton University Press.

Moscadelli, M. (2004). The modeling of operational risk: experience with the analysis of the data collected by the Basel Committee. Working paper, Banca D'Italia. Available at `http://search.ssrn.com/sol3/papers.cfm?abstract_id=557214`.

Neslehova, J., P. Embrechts, and V. Chavez-Demoulin (2006). Infinite mean models and the LDA for operational risk. *Journal of Operational Risk 1*(1), 3–25.

Nystrom, K. and J. Skoglund (2002). Quantitative operational risk management. Swedbank, Group Financial Risk Control. Available at `http://www.gloriamundi.org/picsresources/nsqopm.pdf`.

Operational Riskdata eXchange Association (2007). ORX reporting standards: an ORX member's guide to operational risk event/loss reporting. Available at `http://www.orx.org/`.

Powojowski, M. R., D. Reynolds, and H. J. H. Tuenter (2002). Dependent events and operational risk. *Algorithmics Research Quarterly 4*(2).

Reshetar, A. (2004). Operational risk and the effect of diversification on capital charge. Available at http://www.gloriamundi.org/picsresources/are.pdf.

Reshetar, A. (2008). Dependence of operational losses and the capital at risk. Swiss Banking Institute, University of Zurich. Available at http://www.gloriamundi.org/picsresources/gr_dol.pdf.

Wuthrich, M. V. (2003). Asymptotic value-at-risk estimates for sums of dependent random variables. *Astin Bulletin 33*(1), 75–92.