

LMU

LUDWIG-  
MAXIMILIANS-  
UNIVERSITÄT  
MÜNCHEN

CENTER FOR QUANTITATIVE RISK ANALYSIS  
CEQURA



Stefan Mittnik and Tina Yener

# Estimating Operational Risk Capital for Correlated, Rare Events

Working Paper Number 01, 2009  
Center for Quantitative Risk Analysis (CEQURA)  
Department of Statistics  
University of Munich

<http://www.cequra.uni-muenchen.de>



# Estimating Operational Risk Capital for Correlated, Rare Events\*

Stefan Mittnik<sup>a</sup>     Tina Yener<sup>b</sup>

<sup>a</sup> *Center for Quantitative Risk Analysis (CEQURA) and Department of Statistics,*

*University of Munich, Akademiestr. 1/I, 80799 Munich, Germany*

*finmetrics@stat.uni-muenchen.de, +49 (0) 89 21 80 32 24*

*Center for Financial Studies, Frankfurt*

*Stefan Mittnik, PhD is Professor of Financial Econometrics at the University of Munich.*

<sup>b</sup> *Center for Quantitative Risk Analysis (CEQURA) and Department of Statistics,*

*University of Munich, Akademiestr. 1/I, 80799 Munich, Germany*

*tina.yener@stat.uni-muenchen.de, +49 (0) 89 21 80 33 34*

*Tina Yener is a PhD candidate at the Institute of Financial Econometrics at the University of Munich.*

October 2009

---

\*We thank Nicole Bäuerle, Francesca Biagini, Arthur Charpentier, Paul Embrechts, Christoph Hartz, and Sandra Paterlini as well as the participants of the Actuarial and Financial Mathematics Conference in 2008 and the International Workshop on Rare Event Simulation in 2008 for their helpful comments and discussions. We also thank two anonymous referees for their constructive comments which helped us to greatly improve the paper.

## Abstract

We show that the use of conventional correlations for modeling dependencies may lead to a counterintuitive behavior of risk measures, such as Value-at-Risk and Expected Shortfall, in simulation-based assessments of the risk of very rare events. The effect can be avoided in the case of Expected Shortfall by an appropriate design of the simulation setup. This does not hold, however, for the widely used Value-at-Risk measure.

Consequently, the goal of decreasing minimum capital requirements by specifying less-than-perfect correlations, as suggested by the New Basel Capital Accord (*Basel II*), may not be achieved.

**Keywords:** Operational risk, latent variables, mixture models, correlation, rare events

# 1 Introduction

Since the initiation of the New Basel Capital Accord (*Basel II*) in 1999, when operational risk was introduced to the regulatory landscape, the attention to this risk type has risen substantially. The Basel Committee on Banking Supervision (Basel Committee, 2006) defines operational risk as “risk of loss resulting from inadequate or failed internal processes, people and systems or from external events.” The fact that events like bookkeeping errors and terrorist attacks are covered by this characterization illustrates the broad range of risks relative to credit or market risk. Taking this heterogeneity of loss events into account, the Basel Committee categorizes losses into seven event types and eight business lines. Banks are supposed to calculate risk measures for each of these  $7 \times 8 = 56$  combinations, such as “Internal Fraud” in “Trading and Sales” or “Damage to Physical Assets” in “Commercial Banking”.

The risk measure specified by the Basel Committee is the *Unexpected Loss* at a confidence level of 99.9%. Generally speaking, this refers to the 99.9% quantile of the loss distribution (possibly reduced by the *Expected Loss*, that is, the mean of the distribution), commonly known as the 99.9% *Value-at-Risk* (VaR). It measures the maximum loss that will not be exceeded with the specified confidence level and is a widely used risk measure since the 1990s. The total required risk capital under the *Advanced Measurement Approaches* (AMA) is obtained by summing over all 56 event-type/business-line VaRs, a strategy implicitly expecting the joint occurrence of all loss types involved or, in other words, assuming perfect positive correlation between all loss processes. To allow for non-perfect correlations, the Basel Committee permits a bank “. . . to use internally determined correlations [...] provided it can demonstrate to the satisfaction of the national supervisor that its systems for determining correlations are sound, implemented with integrity, and take into account the uncertainty surrounding any such correlation estimates (particularly in periods of stress).” (Basel Committee, 2006, p. 148). Dropping the highly unrealistic assumption of perfect dependence (i.e., summing the Unexpected Losses of all cells) and relying on realistic correlation estimates should decrease the calculated risk capital. Therefore, banks should have a strong interest in developing and establishing adequate assessment approaches. This expected decrease in estimated risk capital due to less than perfect correlations of the loss processes is the focus here. Specifically, we investigate whether a general rule can be established about risk-capital requirements and less than perfect correlations. Second, we analyze how the model specification affects such a rule.

The *Loss Distribution Approach* (LDA) is by far the most prominent among the AMA, relying on techniques well-known from actuarial applications (Klugman et al., 2004). In its standard form, loss distributions are modeled at the event-type/business-line level,

and the resulting VaR figures are added up. However, dependencies can be introduced between cells, typically by assuming dependent frequencies of occurrence. As an example of such dependence between frequencies, one could think of storm losses (i.e., event type “Damage to Physical Assets”) which typically occur clustered during certain seasons of the year. At the same time, these damages can cause events of the type “Business Disruption and System Failures”, but also affect several business lines located closely to each other. In this paper, we focus on rare-event losses—such as natural catastrophes or terrorist attacks, rather than “everyday” losses such as common bookkeeping errors—and consider risk-generating processes common in credit-risk analysis (see, e.g., Frey and McNeil, 2002, 2003). However, we assume broader parameter ranges than those typically adopted. We confine ourselves to analyzing the frequency part of operational losses to check for the impact of dependent occurrences and disregard the severity dimension by assuming fixed severities. Therefore, in our notion, “risk” is represented by the number of event occurrences.

The main finding of our paper consists of the observation that the assumption of a less than perfect correlation between loss processes does not necessarily imply a decrease in risk capital. Specifically, for all distributional assumptions considered, there may arise situations where a decrease in correlation leads to an *increase* in risk capital estimates, thus reversing the desired effect.

Since the work of Artzner et al. (1999), it is well-known that VaR is not a coherent risk measure due to the lack of subadditivity. In the operational risk context, this means that the joint risk—if measured by VaR—of two event-type/business-line cells may exceed the sum of the individual risks measured for the two single cells. The *Expected Shortfall* (ES) measure, which does fulfil the subadditivity criterion, is typically recommended as an alternative and therefore also considered within the scope of our simulation study. However, it should be pointed out that the subadditivity property is not being analyzed here. At no point in our simulations, risk measures are added up; they are rather calculated based on a sum of *losses*, generated under different correlation assumptions.

The paper is organized as follows. Section 2 introduces the general modeling framework as well as our concrete approach to including dependencies via latent variables. The setup and the results of the simulations are presented in Section 3. Section 4 discusses consequences of our findings and Section 5 concludes.

## 2 Modeling Dependent Operational Losses

In the operational risk modeling framework, the total aggregate loss,  $L$ , is computed by

$$L = \sum_{j=1}^J L_j = \sum_{j=1}^J \sum_{i=1}^{N_j} X_{i,j}, \quad (1)$$

where  $J$  denotes the number of event–type/business–line combinations (with  $J = 56$  using first–level definitions), and  $L_j$  refers to the loss associated with cell  $j$ , which is simply given by the sum of the severities of the individual losses  $X_{i,j}$  with  $i = 1, \dots, N_j$ . There is a number of ways of introducing dependencies in (1).

Dependencies among frequencies and/or severities can exist *within* or *between* cells. The latter is an important focus of research on operational risk. The reason is that the standard LDA assumes that for a given cell  $j$ , quantities  $X_1, X_2, \dots, X_N$ , and  $N$  are independent and that the  $X_i$  are independent and identically distributed random variables.

Analogously, loss dependencies between cells can be introduced by allowing for dependencies among frequencies  $N_j$ ,  $j = 1, \dots, J$ , among severities,  $X_j$ , or by directly specifying dependencies among aggregate cell losses,  $L_j$ . Pfeifer and Nešlehová (2004) and Chavez-Demoulin et al. (2006) construct dependent point processes for the frequency variable,  $N$ . Aue and Kalkbrener (2006) model frequency dependence via a Gaussian copula.<sup>1</sup> Dependent frequencies can also be obtained using common shock models, as, for example, in Lindskog and McNeil (2003). A crucial problem when working with dependent count data is that the dependence structure is no longer solely characterized by the copula. This follows from the non–uniqueness of copulas for discrete data. Though we can simulate dependent frequencies using copulas, drawing inference for the copula parameter from multivariate count data is not at all straightforward (see Genest and Nešlehová, 2007). Specifying dependence among severities has been favored, for example, by Chapelle et al. (2004) and Reshetar (2008). Here, a conceptual problem arises, as dependent severities contradict the assumptions of the standard LDA (see, for example, Frachot et al., 2004, on this issue). The assumption that severities in cells  $i$  and  $\ell$  are correlated (i.e.,  $X_{i,j}$  and  $X_{\ell,k}$  are correlated for all  $i = 1, \dots, N_j$  and  $\ell = 1, \dots, N_k$ ) imposes correlation between severities *within* cells. This contradicts the standard LDA assumption.

In view of the difficulties of “bottom–up” approaches which model in terms of individual loss frequencies and severities, dependence is often introduced directly at the level of aggregate cell losses. Examples are Nyström and Skoglund (2002), Di Clemente and Romano (2003) and Dalla Valle et al. (2008).

---

<sup>1</sup>See Bee (2005) for an analysis of the Gaussian copula’s effect on risk capital estimates.

In this study, to not divert from the focus and to keep the setup as simple as possible, we concentrate exclusively on frequency dependence. That is, we do not explicitly model loss severities, and thereby, adhere to the standard LDA.<sup>2</sup> By doing so, our analysis reduces to modeling the number of loss occurrences. We consider the generic loss variable

$$L = \sum_{i=1}^n Y_i ,$$

composed of  $n$  individual loss processes,  $Y_i$ ,  $i = 1, \dots, n$ . Loss  $Y_i$  is assumed to be either a Bernoulli (0–1) random variable indicating an event occurrence or Poisson distributed, counting the number of event occurrences for process  $i$ .

This setup allows to introduce loss dependence in a variety of ways. For example,  $n = 56$  could represent the number of business–line/event–type combinations, so that  $L$  denotes the total aggregate loss of an institution. On the other hand,  $n$  could be the number of technical components for one business–line/event–type combination. Then,  $L$  represents the losses for one of the 56 cells, assuming within–cell dependencies and, thus, leaving the standard LDA. Intermediate setups, where  $n$  denotes the number of potential failures (such as technical components or transactions) for several cells, are also possible. Whatever the interpretation of  $n$ , risk capital will always be estimated from the distribution of  $L$ . In contrast to Embrechts and Puccetti (2008), who consider the aggregation of risk capital estimates, we assume that risk capital is always estimated from the joint loss distribution, implying that losses in different cells are modeled jointly. If  $n = 56$  refers to all cells, this amounts to the “top–down approach” of Embrechts and Puccetti (2008).

Except in special cases, the cumulative distribution function of  $L$ , denoted by  $F_L$ , is not available in closed form. Therefore, risk capital estimates based on quantiles or expected values cannot be derived analytically but rather have to be obtained via Monte–Carlo simulations. To do so, we simulate  $Y$ –processes and induce dependencies between  $Y_i$  and  $Y_j$ ,  $i \neq j$ , by introducing latent variables.

## 2.1 Latent–variable models

The idea common to all latent–variable specifications is that there exists a second layer of—possibly observable—latent variables which drive the discrete counting process for the observed loss occurrences. Formally, a latent–variable model (LVM) can be defined as follows, cf. McNeil et al. (2005).

---

<sup>2</sup>Dependent severities could easily be handled by introducing an additional simulation step, drawing a loss amount per occurrence. An obvious candidate for a severity distribution is, for example, the multivariate lognormal distribution with the correlation matrix reflecting the dependence structure among severities.

**Definition** (*Latent-variable Model*) Let  $X = (X_1, \dots, X_n)'$  be a random vector and  $D \in \mathbb{R}^{n \times m}$  a deterministic matrix with elements  $d_{ij}$ . Suppose that

$$S_i = j \Leftrightarrow d_{ij} < X_i < d_{i,j+1}, \quad i \in \{1, \dots, n\}, j \in \{0, \dots, m\},$$

where  $d_{i0} = -\infty$ ,  $d_{i,m+1} = \infty$ . Then,  $(X, D)$  is a *latent-variable model* for the state vector  $S = (S_1, \dots, S_m)'$ , where  $X_i$  are the latent variables and  $d_{ij}$  the thresholds of the LVM.

For our purposes, we introduce variable  $Y_i$ , defined by

$$Y_i = 1 \Leftrightarrow S_i = 0 \quad \text{and} \quad Y_i = 0 \Leftrightarrow S_i > 0,$$

to indicate the “occurrence” and “non-occurrence” of an event, i.e., we only distinguish between these two states. The probability of occurrence for process  $i$ ,  $\pi_i$ , is defined by

$$\Pr[Y_i = 1] = \Pr[X_i \leq d_{i1}] = \pi_i.$$

In the credit-risk literature,  $Y_i = 1$  indicates a “default” of counterparty  $i$ , meaning that obligor  $i$  cannot make its payments. In structural credit-risk models, the latent variable is interpreted as the obligor’s assets. If their value falls below some threshold (the default boundary) the obligor defaults. However, as pointed out by Duffie et al. (2009), credit risk portfolio losses are not driven by one common factor, and neglecting unobserved latent variables may lead to an underestimation of risk.

This concept of latent variables can be adapted to suit operational risk settings, as is done, for example, in Peters et al. (2009). Here, the underlying variables can be interpreted as *Key Risk Drivers* (KRDs) (Nyström and Skoglund, 2002); that is, they induce indirect dependence by driving several loss processes. The interpretation of a KRD will typically depend on the event type (and/or business line) considered; while it may represent the predisposition for natural disasters in the “Damage to Physical Assets” event type, a KRD for “Internal Fraud” will surely contain the quality of internal controls. This illustrates that one cannot, in general, state whether KRDs are exogenous to the institution or whether they can be controlled by the risk manager. When modeling operational risk, it may be reasonable to set  $Y_i \in \{0, 1, 2, \dots\}$ , representing the *number* of loss events (rather than the two outcomes “default” or “no default”), in which case the Poisson distribution rather than the Bernoulli distribution is appropriate. The Poisson distribution is a natural candidate since it approximates sums of Bernoulli random variables with low success probabilities. This property will be utilized in the mixture model representation below; however, our findings apply both to the Bernoulli and the Poisson formulation.

We want to construct a setup in which the probability of the occurrence of an event can depend on events in other processes, as, for example, a system breakdown in one corporate

division may propagate to others causing a failure there. In LVMs, this is modeled by allowing for dependencies among the latent variables, inducing dependencies among event occurrences in an indirect fashion. Restricting ourselves to linear dependence, we distinguish between *latent correlation* among the  $X_i$  and *observed correlation* among the  $Y_i$ . The latter is given by (see, e.g. McNeil et al., 2005, p. 344)

$$\rho_Y = \frac{\text{Cov}[Y_i, Y_j]}{\sqrt{\text{Var}[Y_i] \cdot \text{Var}[Y_j]}} = \frac{\text{E}[Y_i Y_j] - \pi_i \pi_j}{\sqrt{\pi_i(1 - \pi_i)\pi_j(1 - \pi_j)}}, \quad (2)$$

where  $\text{E}[Y_i Y_j] = \Pr[Y_i = 1, Y_j = 1] = \Pr[X_i \leq d_{i1}, X_j \leq d_{j1}]$  denotes the joint cumulative distribution function of the latent variables associated with processes  $i$  and  $j$ . The observed correlation,  $\rho_Y$ , is called “default correlation” in the credit–risk literature, as opposed to the (latent) “asset correlation”,  $\rho_X$ , which refers to the linear dependence between latent variables. From (2) it follows that observed correlations depend on marginal occurrence probabilities,  $\pi_i$  and  $\pi_j$ , and on latent correlation,  $\rho_X$ , the latter entering via  $\text{E}[Y_i Y_j]$ . Relationship (2) has been studied, for example, by Gersbach and Lipponer (2003) and Foulcher et al. (2005).

Normal–variance mixtures are obvious and widely used candidates for the distribution of latent variables. In normal–variance mixtures, latent variables can be written as

$$X = \mu + \sqrt{W}Z,$$

where  $Z \sim \text{N}(0, \Sigma)$ ,  $W$  is a scalar random variable independent of  $Z$ , and  $\mu$  is a constant. An event occurs for process  $i$  if  $X_i \leq d_{i1}$  or

$$Z_i \leq \frac{d_{i1} - \mu}{\sqrt{W}}.$$

## 2.2 Mixture models

Mixture models can arise when distributional parameters do not remain constant. It seems plausible that, for example, in periods when tectonic plates move, the probability of earthquakes rises, that storms are more likely to happen in one season than in others, or that a management change in a company can affect the probability of business disruptions. Therefore, in an operational–risk context, it seems to be a realistic assumption that distributional parameters are subject to changes, i.e., that they themselves are random. A formal definition of a special mixture model in the spirit of McNeil et al. (2005) is as follows.

**Definition** (*Bernoulli Mixture Model*) Let  $Y = (Y_1, \dots, Y_n)'$  be a random vector in  $\{0, 1\}^n$  and  $\Psi = (\Psi_1, \dots, \Psi_p)'$ ,  $p < n$ , be a factor vector. Then,  $Y$  follows a *Bernoulli mixture model* with factor vector  $\Psi$  if there exist functions  $p_i : \mathbb{R}^p \rightarrow [0, 1]$  such that conditional on  $\Psi$  the elements of  $Y$  are independent Bernoulli random variables with  $\Pr[Y_i = 1 | \Psi = \psi] = p_i(\psi)$ .

It is also possible to define  $Y$  to be conditionally Poisson distributed. Then,  $Y$  is a count variable rather than a binary variable, giving rise to a *Poisson mixture model*. Both models can be transformed into each other by setting  $Y = \mathcal{I}_{(0,\infty)}(\tilde{Y})$ , where  $\mathcal{I}$  denotes the indicator variable and  $\tilde{Y} \sim \text{Poi}(\lambda)$ . The parameters are related via  $p_i = 1 - e^{-\lambda_i}$ , a property which we use to simulate from both models in a comparable way. To keep the setup simple, we examine only exchangeable mixture models, where conditional probabilities of event occurrence are identical, i.e.,  $p_i(\psi) = p(\psi)$ . Defining the new random variable  $Q = p(\Psi)$ , the observed correlation between indicator variables can then be obtained from

$$\rho_Y = \frac{\pi_2 - \pi^2}{\pi - \pi^2},$$

where  $\pi = E[Q]$  and  $\pi_k = E[Q^k]$ , the latter resulting from conditional independence.<sup>3</sup>

In fact, LVMs and Bernoulli mixture models can be viewed as two different representations of the same underlying mechanism. The following lemma from Frey and McNeil (2003) states the condition for such an equivalence.

**Lemma 1** *Let  $(X, D)$  be an LVM with  $n$ -dimensional random vector  $X$ . If  $X$  has a  $p$ -dimensional conditional independence structure with conditioning variable  $\Psi$ , the default indicators  $Y_i = \mathcal{I}_{X_i \leq d_{i1}}$  follow a Bernoulli mixture model with conditional event probabilities  $p_i(\psi) = \text{Pr}[X_i \leq d_{i1} | \Psi = \psi]$ .*

The LVM  $(X, D)$ , where  $X$  is a normal–variance mixture and  $Z$  is governed by a one–factor structure, can be written as

$$X_i = \sqrt{W} Z_i, \tag{3a}$$

$$Z_i = \sqrt{\rho_X} \Psi + \sqrt{1 - \rho_X} \varepsilon_i, \tag{3b}$$

where  $\rho_X$  is the latent correlation,  $\varepsilon_i \stackrel{\text{iid}}{\sim} N(0, 1)$ , and  $\Psi \sim N(0, 1)$  is the only factor and conditioning variable. We thus obtain a conditional independence structure for  $X$  which allows us to proceed using the equivalent mixture model representation.

Below, we do not only want to analyze the effect of correlation on risk–capital estimates for rare events, but also to assess consequences in terms of model risk, i.e., the use of a model which does not accurately describe the processes in place. Therefore, we compare the impact of different distributional assumptions that are frequently adopted in the credit–risk framework. To do this systematically, we calibrate the mixture models with the help of latent and observed correlations. The simplest and most common LVM, serving

---

<sup>3</sup>Due to conditional independence, one obtains

$$\pi_k = \text{Pr}[Y_1 = 1, \dots, Y_k = 1] = E[E[Y_1 \cdots Y_k | Q]] = E[E[Y_1 | Q] \cdots E[Y_k | Q]] = E[Q^k].$$

here as benchmark model, assumes multivariate normal latent variables. In this case, the distribution of  $X_i$  in (3a) corresponds to that of  $Z_i$ , implying that  $W = 1$ . Standardizing  $X_i$  and inserting the threshold expression  $d_i = \Phi^{-1}(\pi)$ , one obtains the conditional default probability

$$p(\psi) = \Pr[X_i \leq d_i | \Psi = \psi] = \Phi \left( \frac{\Phi^{-1}(\pi) - \sqrt{\rho_X} \psi}{\sqrt{1 - \rho_X}} \right). \quad (4)$$

In order to assess model risk, we adapt modifications suggested and analyzed in the context of credit portfolios (cf. Frey and McNeil, 2002, 2003). The first generalization of our benchmark model allows for tail dependence and a fat-tailed multivariate distribution for the latent variables by assuming a multivariate Student- $t$  distribution. In this case,  $\rho_X = 0$  means that the latent variables are uncorrelated, but they are no longer independent. The degree-of-freedom parameter of the  $t$  distribution,  $\nu$ , adds flexibility and allows to control the degree of fat-tailedness of the latent variables. The setup given in (3) remains valid, but for  $X_i$  to have a Student- $t$  distribution with  $\nu$  degrees of freedom, the normally distributed random variable  $Z_i$  has to be multiplied by an inverse-gamma distributed random variable; i.e.,  $W \sim \text{InvGam}(\nu/2, \nu/2)$ . The resulting conditional occurrence-probability is

$$p(\psi) = \Pr[X_i \leq d_i | \Psi = \psi] = \Phi \left( \frac{t_\nu^{-1}(\pi) W^{-1/2} - \sqrt{\rho_X} \psi}{\sqrt{1 - \rho_X}} \right), \quad (5)$$

where  $t_\nu^{-1}(\cdot)$  denotes the inverse of the  $t$  distribution with  $\nu$  degrees of freedom.

Leaving the framework defined by (3), two additional modifications are considered. First, we assume a Beta distribution for the mixing variable. It is a natural extension, because the interval  $[0, 1]$  is the domain, and therefore allows us to interpret the mixing variable as conditional probability. Moreover, it leads to an analytically tractable model. We therefore consider a mixing variable  $p(\psi) = Q \sim \text{Beta}(a, b)$ . As the moments of a Beta distribution can be directly calculated from the distributional parameters,  $a$  and  $b$ , we can easily derive unconditional occurrence-probabilities<sup>4</sup>,

$$\pi_k = \frac{\beta(a+k, b)}{\beta(a, b)} = \prod_{j=0}^{k-1} \frac{a+j}{a+b+j}, \quad (6)$$

and the observed correlation

$$\rho_Y = \frac{1}{a+b+1}.$$

---

<sup>4</sup>Here, we are making use of the representation of the Beta function,  $\beta(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$ , in terms of the Gamma function,  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ , which is

$$\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Equation (6) then follows from the general recursion  $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ .

Below, these relationships are used for model calibration.

Finally, we also consider an Archimedean copula

$$C(u_1, \dots, u_d) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_d))$$

for the latent variables, where  $\phi$  (the “generator”) refers to the the inverse of the Laplace transform of cumulative distribution function  $G$  on  $\mathbb{R}$ . This choice of copula, as pointed out by Frey and McNeil (2003), is mainly motivated by its simplicity for calibration and simulation. Furthermore, Archimedean copulas allow—in contrast to the Gaussian copula—for tail dependence, the exact form of which depends on the form of the generator. To simulate latent variables with an Archimedean copula, we assume a factor,  $\Psi$ , and a sequence of uniform random variables,  $U_1, \dots, U_d$ , conditionally independent given  $\Psi$ , with

$$\Pr[U_i \leq u | \Psi = \psi] = \exp(-\psi\phi(u)) , \quad u \in [0, 1] .$$

Conditional occurrence probabilities can be calculated from

$$Q = p(\psi) = \Pr[U_i \leq \pi | \Psi = \psi] = \exp(-\psi\phi(\pi)) .$$

In the special case of a Clayton copula, the generator takes the form  $\phi(t) = t^{-\theta} - 1$  and  $\Psi$  needs to be gamma distributed, i.e.,  $\Phi \sim \text{Ga}(1/\theta, 1)$ . The resulting copula is characterized by lower tail dependence.<sup>5</sup> With this, we obtain the bivariate occurrence probability

$$\pi_2 = \phi^{-1}(\phi(\pi) + \phi(\pi)) = (2\pi^{-\theta} - 1)^{-1/\theta} , \quad \theta > 0 , \quad (7)$$

which is used for calibration to the benchmark model.

## 3 Simulation

### 3.1 Simulation Setups

In order to obtain risk–capital estimates, the loss distribution,  $F_L$ , needs to be derived. As this function is not available in closed form, we construct its empirical distribution function

$$\widehat{F}_L(l) = \frac{1}{B} \sum_{b=1}^B \mathcal{I}_{[0,l]}(l_b)$$

---

<sup>5</sup>We thus impose tail dependence in that part of the latent variables’ distribution which is relevant for loss event occurrences. The coefficient of lower tail dependence for the Clayton copula is given by  $\lambda_L = 2^{-1/\theta}$ , while for the upper tail,  $\lambda_U = 0$  holds.

via Monte–Carlo simulation, where  $l_b$  refers to the loss from replication  $b$  and  $B$  replications are used altogether. The resulting  $\text{VaR}_{1-\alpha}$  value, based on these  $B$  replications, is then given by

$$\text{VaR}_{1-\alpha} = \widehat{F}_L^{-1}(1 - \alpha) ,$$

and that for  $\text{ES}_{1-\alpha}$  by

$$\text{ES}_{1-\alpha} = \frac{1}{B} \sum_{b=1}^B l_b \mathcal{I}_{[\text{VaR}, \infty]}(l_b) .$$

In the benchmark model, risk capital estimates are generated as follows:

1. Set values for the observed occurrence probabilities  $\pi$  and latent correlations  $\rho_X$
2. Simulate a (standard normally distributed) factor realization,  $\psi$ , and calculate the conditional occurrence–probability  $p(\psi)$  from (4)
3. Conduct  $n$  Bernoulli trials, using  $p(\psi)$  as success probability, and sum up the number of event occurrences
4. Repeat these steps  $B$  times and calculate VaR and ES for the resulting empirical distribution.

For the distributional specifications considered, this procedure has to be modified accordingly. In order to draw from a Student- $t$  distribution for the latent variables, we fix  $\pi$  and  $\rho_X$  to the same values as in the benchmark model. At the same time, we are free to set a value for the degrees-of-freedom parameter  $\nu$ , a smaller value implying fatter tails for the distribution of the latent variables. In addition to drawing a factor realization from  $\Psi \sim \text{N}(0, 1)$ , we also draw from  $W \sim \text{InvGam}(\nu/2, \nu/2)$  and calculate the conditional occurrence–probability  $p(\psi)$  from (5), which is then again used for Bernoulli/Poisson trials in order to simulate loss event occurrences.

In case of a beta mixing distribution, we calculate the bivariate occurrence probability,  $\pi_2$ , implied by  $\pi$  and  $\rho_X$  in the benchmark model.<sup>6</sup> Parameters  $a$  and  $b$  of the beta distribution can then be derived from  $\pi$  and  $\pi_2$  via (6). Drawing from the Beta distribution with parameters  $a$  and  $b$ , we directly obtain the conditional occurrence–probability,  $p(\psi)$ , which is then used to conduct  $n$  Bernoulli/Poisson trials. Again, the procedure is repeated  $B$  times, and values for VaR and ES are derived.

In case of a Clayton copula of latent variables, we first need to determine the value of parameter  $\theta$  implied by the benchmark–model values of  $\pi$  and  $\rho_X$ . In order to do so, we again calculate the corresponding  $\pi_2$  and derive  $\theta$  from  $\pi$  and  $\pi_2$  using (7). We then draw

---

<sup>6</sup> $\pi_2$  denotes the value of the cumulative distribution function of a bivariate normal distribution with correlation  $\rho_X$  at  $(\Phi^{-1}(\pi), \Phi^{-1}(\pi))$ .

factor realizations from  $\Psi \sim \text{Gam}(1/\theta, 1)$  which are used to obtain conditional occurrence probabilities  $p(\psi)$ . Given these, we again conduct  $n$  Bernoulli/Poisson trials and replicate  $B$  times in order to derive values for VaR and ES.

## 3.2 Simulation Results

For each of the models discussed above we simulate event occurrences and estimate risk capital for different levels of latent correlation,  $\rho_X$ . In doing this, we use the multivariate normal LVM as benchmark model to which we calibrate the other models.

In a first simulation study, we assume  $n = 1,000$  loss processes, to match the setup of the study in Frey and McNeil (2001). Specifying  $B = 100,000$  replications, we observe for all models a counterintuitive behavior of the VaR values when occurrence probabilities are low ( $\pi \leq 0.01$ ): Starting from perfect correlation, VaR values increase as correlations decrease. We consider confidence levels between 95% and 99.9% and find that this effect is more pronounced for lower confidence levels, i.e., when decreasing  $\rho_X$ , we observe it first for  $\text{VaR}_{0.95}$ , then for  $\text{VaR}_{0.99}$ , and only for very low occurrence probabilities ( $\pi \leq 0.0001$ ),  $\text{VaR}_{0.999}$  is also affected. An illustration of this phenomenon is given in Figure 1, which plots the 99% VaR (on a logarithmic scale) as a function of the level of latent correlation and occurrence probability,  $\pi$ .

Figure 1 somewhere here.

For  $\pi = 0.01$ , VaR behaves as expected: It increases in  $\rho_X$  over the entire range of latent correlations. However, for lower levels of  $\pi$ , such as  $\pi = 0.005$ , VaR decreases with increasing correlation above a certain threshold of  $\rho_X$ ; the lower  $\pi$ , the lower this threshold value. This effect is the more pronounced, the fatter the tails of the distribution of latent variables. This is shown in Figure 2, where an intermediate occurrence-probability of  $\pi = 0.001$  is held fixed.

Figure 2 somewhere here.

For  $\nu = 100$ , VaR grows as the latent correlation increases to  $\rho_X \approx 0.5$  and decreases for higher levels. The lower  $\nu$ , the broader the range of  $\rho_X$  for which this peculiar behavior occurs. For  $\nu = 4$ , VaR values decrease over the entire range of latent correlations,  $\rho_X$ .

The results for Poisson mixture models are qualitatively the same, as was to be expected from the low level of occurrence probabilities involved. For ES, using 100,000 replications, we obtain ambiguous results. Figure 3 shows a behavior which corresponds to intuition, i.e., ES increases with correlation.

Figure 3 somewhere here.

However, setting a different seed in the simulation can result in ES-increases as correlation decreases when the occurrence probability is very low ( $\pi \leq 1.0e-005$ ).

The results presented so far are based on 100,000 replications. In order to assess the uncertainty in risk capital estimates, we modify the simulation setup and calculate several risk capital estimates for each combination of  $\rho_X$  and  $\pi$ , increasing the number of replications up to  $B = 10,000,000$ . In order to avoid predominance of risk capital estimates of zero, we now set the number of loss processes to  $n = 100,000$  and concentrate on a confidence level of 99.9% (which is the relevant level in operational-risk applications). Due to the substantial computational burden of this task, we hold the occurrence probability fixed at  $\pi = 1.0e-005$ . To assess convergence and reliability of risk capital estimates, we conduct 250 Monte-Carlo simulations for each  $\rho_X$ . Figures 4 to 9 compare boxplots of these 250 risk capital estimates, where we replace the median by the mean over the 250 estimates for assessing convergence.

Figures 4 to 9 somewhere here.

Figure 4 confirms the findings of our first simulations: Above a certain level of latent correlation ( $\rho_X \approx 0.5$ ), VaR estimates may increase with decreasing latent correlation. The dispersion of the 250 risk capital figures is the highest for medium levels of correlation. As was to be expected, it decreases with an increasing number of replications; however, VaR still increases with decreasing correlations in the upper region of  $\rho_X$ . In view of these observations, we conclude that this counterintuitive behavior is not due to convergence issues and cannot be eliminated by increasing the number of replications. It appears that  $B = 1,000,000$  replications are sufficient to obtain reliable VaR figures. Therefore, we restrict our attention to this  $B$ -value when analyzing VaR behavior in the following.

Figure 5 shows the ES figures resulting from the very same simulations. It becomes evident that the risk-capital estimates may, just like in the VaR case, rise with decreasing correlations; but in the case of ES figures, this is not true for their mean. Therefore, increasing the number of replications to  $B^{\max} = 10,000,000$  substantially reduces the occurrence frequency of ES-increases caused by correlation declines. The restriction to  $B^{\max}$  is due to computational limitations.

Figure 6 illustrates that the results from a Poisson mixture model are qualitatively identical to those of the Bernoulli version—except for ES values at very high correlation levels. This is due to the low occurrence probability, making multiple loss occurrences extremely rare. The results from the previous study regarding the fat-tailedness of the latent variables' distribution are confirmed by Figure 7. Just as before, VaR figures decrease in  $\rho_X$  over the entire range of latent correlations. In addition, we observe that the dispersion

of the ES-values is much higher than in the case of normally distributed latent variables, implying that ES-increases caused by correlation declines can be observed more often. Figures 8 and 9 confirm the previous findings for the Beta mixing distribution and the Clayton copula of latent variables.

## 4 Interpretation of Results

The aim of our simulation study was to find out whether there is some general rule for the effects of (less than perfect) correlation between operational loss processes. We introduce dependence through latent variables and assume that the event-type/business-line cells under consideration are modeled jointly before estimating risk capital; i.e., we do not assess subadditivity properties.

Surprisingly, we find that for all distributional assumptions considered, risk capital estimates do not necessarily decrease when departing from the assumption of perfect positive correlation. This counterintuitive effect is observed for low occurrence probabilities ( $\pi \leq 0.01$ ) and only above a certain threshold for  $\rho_X$ . The threshold decreases as the tail dependence of the latent variables increases. The effect is of practical relevance only for rather “extreme” situations, namely, for risk capital estimation at very high confidence levels for very rare, but (possibly tail-)dependent events. We argue that—beyond the 99.9% confidence level prescribed by the Basel Committee—this setup is exactly what can be expected for certain event types. Examples are the joint modeling of earthquake losses across business lines located close to each other, or internal fraud cases in business lines subjected to the same control environment. Both cases involve events, which rarely occur. But, in case they do, they will probably hit several cells simultaneously. We have to keep in mind that the correlation threshold above which this effect arises is that of the *latent* correlation. The implied *observed* correlation, e.g., between earthquake losses, can be much lower.

Our findings reveal that the idea of generating incentives for modeling dependencies more realistically by reducing capital requirements may just cause the opposite and increase capital requirements.

The explanation for the unexpected behavior of the risk capital estimate with respect to different correlation assumptions is illustrated in Figure 10, showing each 10,000 draws from a bivariate normal distribution under two correlation assumptions. The solid lines represent the thresholds implied by the occurrence probability of  $\pi = 0.01$ . In the upper plot, with latent correlation  $\rho_X = 0.1$ , this threshold leads to four joint “occurrences”

(in the southwest quadrant) and 9,798 joint “non-occurrences”. In the lower plot, with correlation  $\rho_X = 0.9$ , the concentration on extremes leads to 94 joint “occurrences” and 9,854 joint “non-occurrences”. Hence, high correlation not only leads to more joint events, but also to more joint “non-events”. And it is the increase in joint “non-events” which lowers the Value-at-Risk estimate as correlation rises.

Figure 10 somewhere here.

Note that we kept the setup of our simulation study as simple as possible. That is, we assumed exchangeable latent variables as well as constant latent correlations and occurrence probabilities. In practice, these assumptions will not be met, as, for example, occurrence probabilities may change over time. This may require a dynamic treatment. Note that incorporating loss severities—which we neglect in our analysis to focus on the particular problem under investigation—is easily accomplished by simply adding another simulation layer. As we tend to deal with low-frequency/high-severity event types, the severity dimension may be the dominant aspect in such simulations. These issues will be subject to further research.

## 5 Conclusion

Introducing less than perfect dependencies should lead to a more realistic description of potential loss events. Our results show that it is important to assess the impact of correlations given the chosen modeling framework. In the case of rare events, simulated values for risk measures, such as Value-at-Risk and Expected Shortfall, can increase as the level of correlation decreases. The parameter ranges for which this phenomenon occurs may be less relevant for credit risk, but they arise in operational risk applications, where, for example, several business lines at close locations are likely to be affected by the same catastrophic event.

For Expected Shortfall, this problem can be circumvented by choosing an appropriate design of the Monte-Carlo setup. Unfortunately, this is not so for the widely used Value-at-Risk, which systematically declines above certain levels of latent correlations. The extent of the problem depends on the observed occurrence probabilities, the confidence level and the fat-tailedness of the distribution of the latent variables. If the clustering of realizations at zero (“joint non-occurrences”) is not in line with the true risk-generation mechanisms, risk capital may be severely underestimated. In this case, other statistical concepts of dependency should be considered for risk-capital calculation.

A practical implication of our analysis is that the inclusion of non-perfect correlations in models used for assessing minimum capital requirements for operational risk may, in fact,

increase the computed requirements.

## References

- Artzner, P., Delbaen, F., Eber, J. M., Heath, D., 1999. Coherent Measures of Risk. *Mathematical Finance* 9 (3), 203–228.
- Aue, F., Kalkbrenner, M., 2006. LDA at Work: Deutsche Bank’s Approach to Quantifying Operational Risk. *Journal of Operational Risk* 1 (4), 49–93.
- Basel Committee, June 2006. International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Technical Report, Bank for International Settlements, Comprehensive Version.
- Bee, M., 2005. Copula-Based Multivariate Models with Applications to Risk Management and Insurance. Working paper, Università degli Studi di Trento - Faculty of Economics.
- Chapelle, A., Crama, Y., Hübner, G., Peters, J.-P., 2004. Basel II and Operational Risk: Implications for risk measurement and management in the financial sector. Working Paper 51, National Bank of Belgium.
- Chavez-Demoulin, V., Embrechts, P., Nešlehová, J., 2006. Quantitative Models for Operational Risk: Extremes, Dependence and Aggregation. *Journal of Banking and Finance* 30, 2635–2658.
- Dalla Valle, L., Fantazzini, D., Giudici, P., 2008. Copulae and operational risks. *International Journal of Risk Assessment and Management* 9 (3), 238–257, available at SSRN: <http://ssrn.com/abstract=944174>.
- Di Clemente, A., Romano, C., 2003. A Copula-Extreme Value Theory Approach for Modelling Operational Risk. In: *Operational Risk Modelling and analysis: Theory and Practice*. Risk Books.
- Duffie, D., Eckner, A., Horel, G., Saita, L., 2009. Frailty correlated default. *The Journal of Finance* 64 (5), 2089–2123.
- Embrechts, P., Puccetti, G., 2008. Aggregating operational risk across matrix structured loss data. *The Journal of Operational Risk* 3 (2), 29–44.
- Foulcher, S., Gouriéroux, C., Tiomo, A., 2005. Latent Variable Approach to Modelling Dependence of Credit Risks: Application to French Firms and Implications for Regulatory Capital. Working Paper CREF 05-01, HEC, Montreal.

- Frachot, A., Roncalli, T., Salomon, E., 2004. The Correlation Problem in Operational Risk. Working Paper , Crédit Lyonnais.
- Frey, R., McNeil, A. J., 2001. Modelling Dependent Defaults. Preprint, Universität and ETH Zürich.
- Frey, R., McNeil, A. J., 2002. VaR and expected shortfall in portfolios of dependent credit risks: Conceptual and practical insights. *Journal of Banking & Finance* 26, 1317–1334.
- Frey, R., McNeil, A. J., 2003. Dependent Defaults in Models of Portfolio Credit Risk. *Journal of Risk* 6 (1), 59–92.
- Genest, C., Nešlehová, J., 2007. A primer on copulas for count data. *The ASTIN Bulletin* 37, 475–515.
- Gersbach, H., Lipponer, A., 2003. Firm Defaults and the Correlation Effect. *European Financial Management* 9 (3), 361–378.
- Klugman, S. A., Panjer, H. H., Willmot, G., 2004. *Loss Models: From Data to Decisions*. Wiley.
- Lindskog, F., McNeil, A., 2003. Common Poisson Shock Models: Applications to Insurance and Credit Risk Modelling. *ASTIN Bulletin* 33 (2), 209–238.
- McNeil, A., Frey, R., Embrechts, P., 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- Nyström, K., Skoglund, J., 2002. Quantitative Operational Risk Management. Working paper, Swedbank.
- Peters, G. W., Shevchenko, P. V., Wüthrich, M. V., 2009. Dynamic operational risk: modeling dependence and combining different sources of information. *The Journal of Operational Risk* 4 (2), 69–104.
- Pfeifer, D., Nešlehová, J., 2004. Modeling and generating dependent risk processes for IRM and DFA. *ASTIN Bulletin* 34 (2), 333–360.
- Reshetar, G., 2008. Dependence of operational losses and the capital at risk. Tech. rep., University of Zurich—Swiss Banking Institute (ISB).

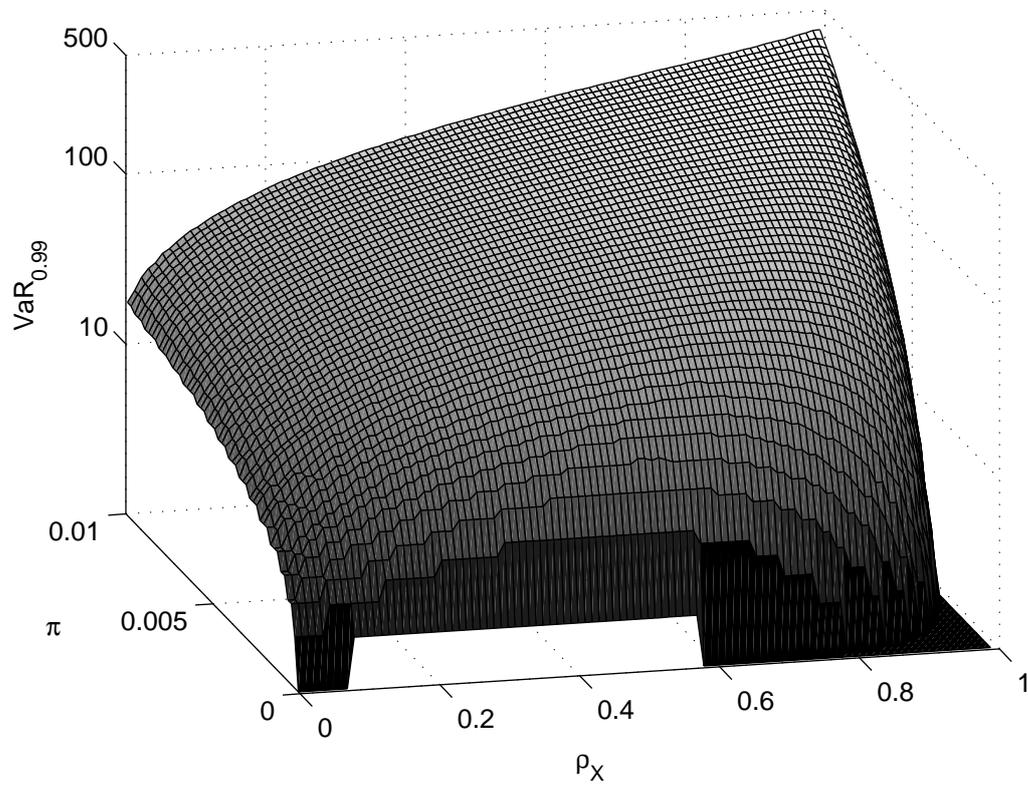


Figure 1: Simulated  $\text{VaR}_{0.99}$  figures from a Bernoulli mixture model with multivariate normal latent variables,  $\pi \in [1.0\text{e-}003, 0.01]$

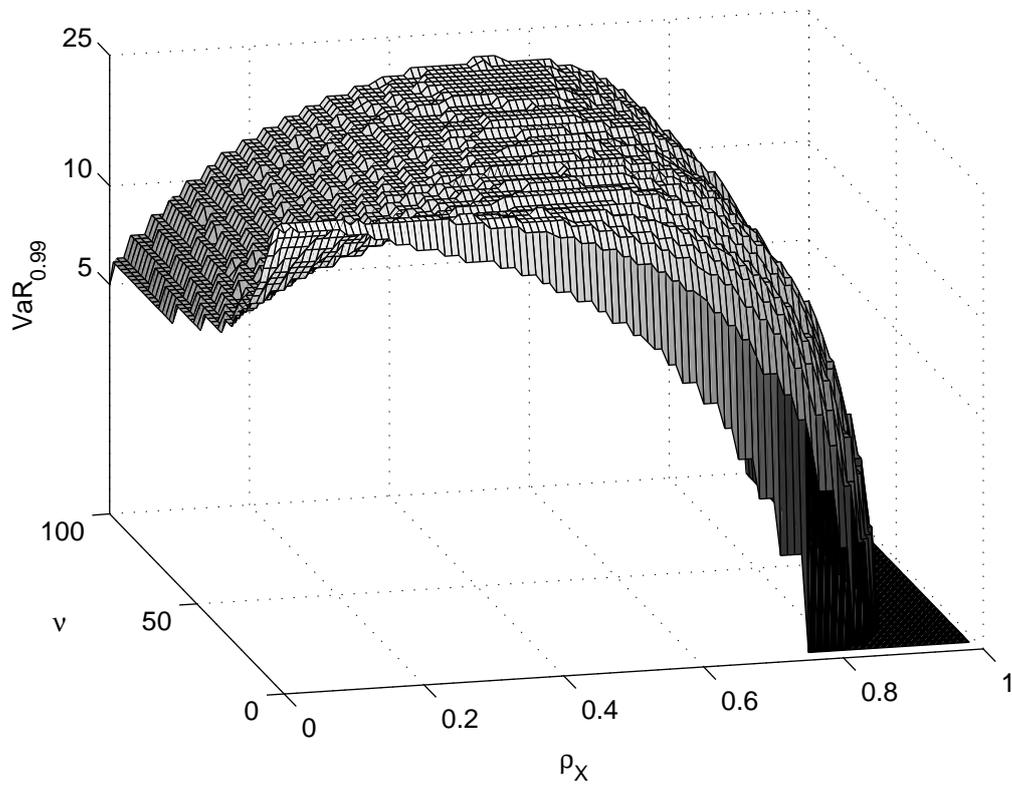


Figure 2: Simulated  $\text{VaR}_{0.99}$  figures from a Bernoulli mixture model with multivariate  $t$ -distributed latent variables,  $\pi = 0.001$ ,  $\nu \in [4, 100]$

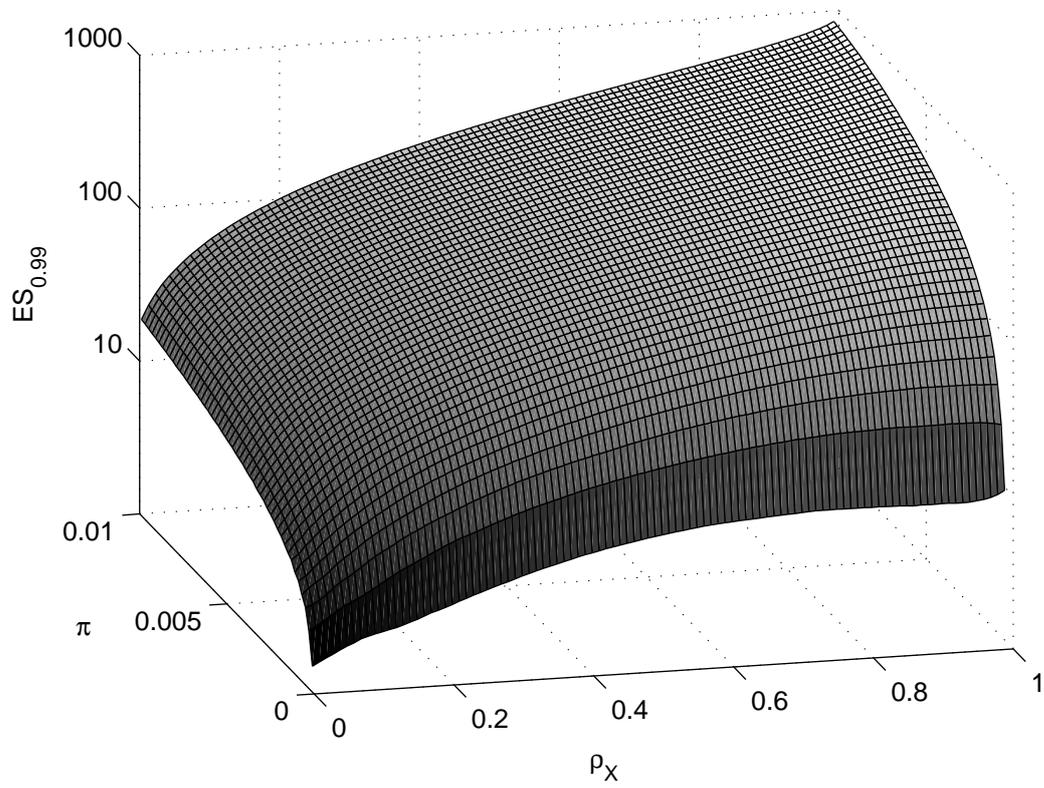


Figure 3: Simulated  $ES_{0.99}$  figures from a Bernoulli mixture model with multivariate normal latent variables,  $\pi \in [1.0e-003, 0.01]$

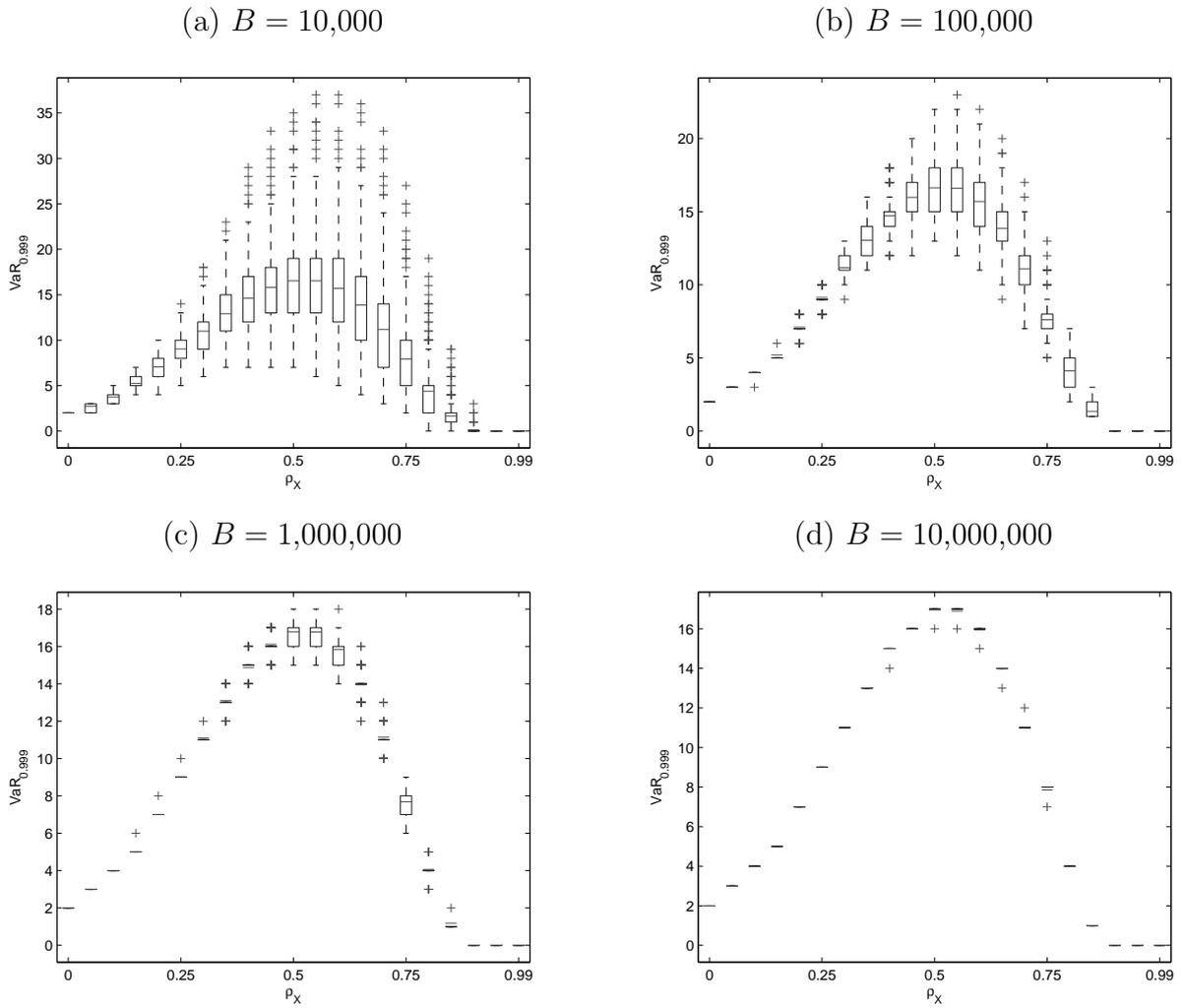
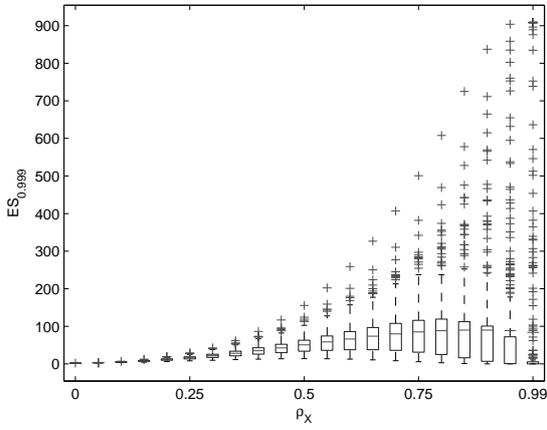
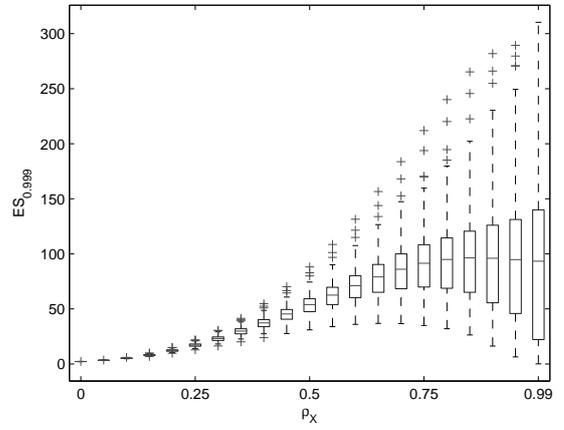


Figure 4: Mean–boxplots of 250 simulated  $\text{VaR}_{0.999}$  figures from a Bernoulli mixture model with multivariate normal latent variables,  $\pi = 1.0\text{e-}005$ ,  $B$  replications

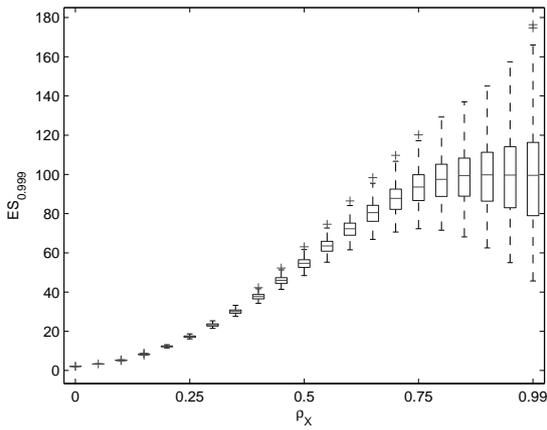
(a)  $B = 10,000$



(b)  $B = 100,000$



(c)  $B = 1,000,000$



(d)  $B = 10,000,000$

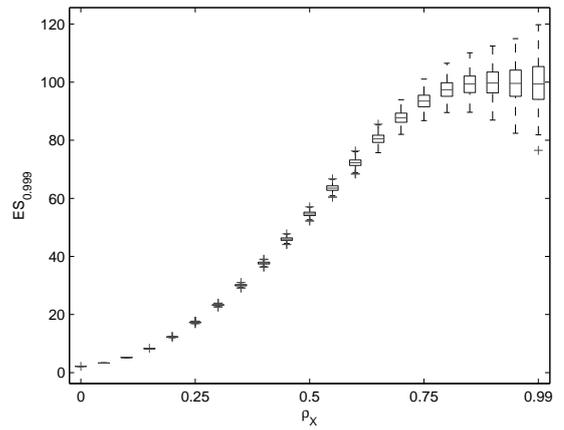
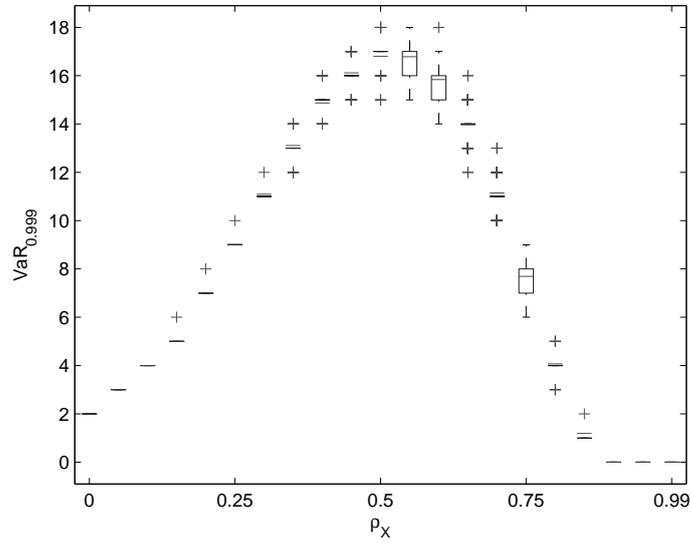


Figure 5: Mean–boxplots of 250 simulated  $ES_{0.999}$  figures from a Bernoulli mixture model with multivariate normal latent variables,  $\pi = 1.0e-005$ ,  $B$  replications

(a)  $\text{VaR}_{0.999}$ ,  $B = 1,000,000$



(b)  $\text{ES}_{0.999}$ ,  $B = 10,000,000$

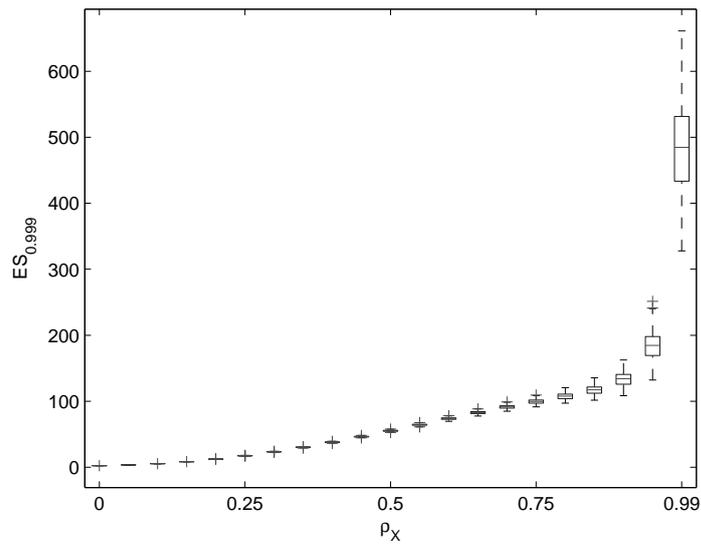
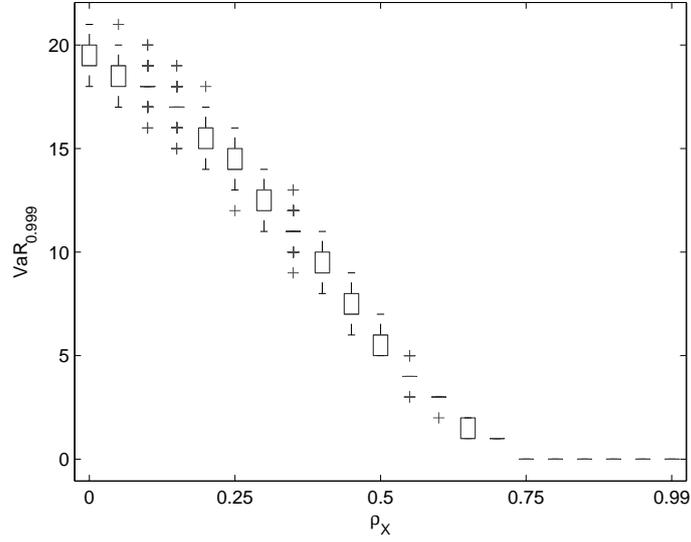


Figure 6: Mean-boxplots of 250 simulated  $\text{VaR}_{0.999}$  and  $\text{ES}_{0.999}$  figures from a Poisson mixture model with multivariate normal latent variables,  $\pi = 1.0\text{e-}005$ ,  $B$  replications

(a)  $\text{VaR}_{0.999}$ ,  $B = 1,000,000$



(b)  $\text{ES}_{0.999}$ ,  $B = 10,000,000$

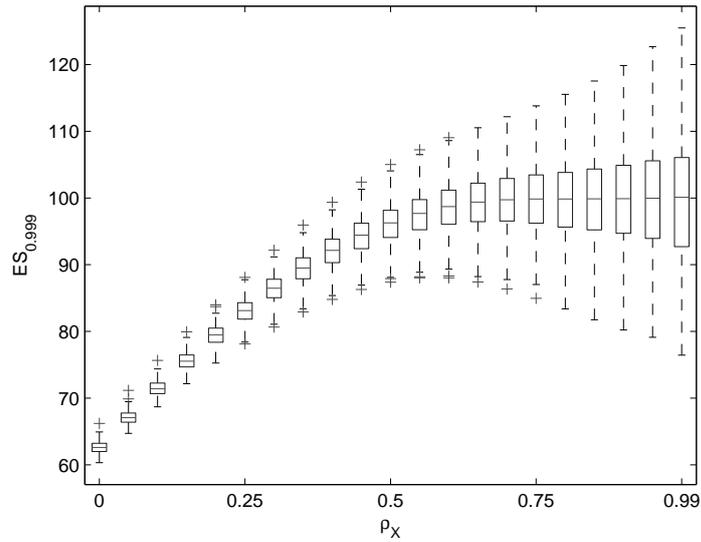
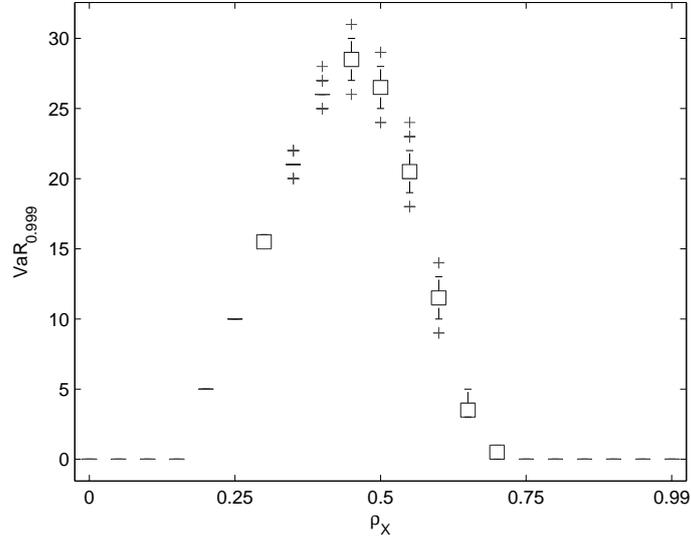


Figure 7: Mean–boxplots of 250 simulated  $\text{VaR}_{0.999}$  and  $\text{ES}_{0.999}$  figures from a Bernoulli mixture model with multivariate  $t$ -distributed latent variables,  $\nu = 4$ ,  $\pi = 1.0\text{e-}005$ ,  $B$  replications

(a)  $\text{VaR}_{0.999}$ ,  $B = 1,000,000$



(b)  $\text{ES}_{0.999}$ ,  $B = 10,000,000$

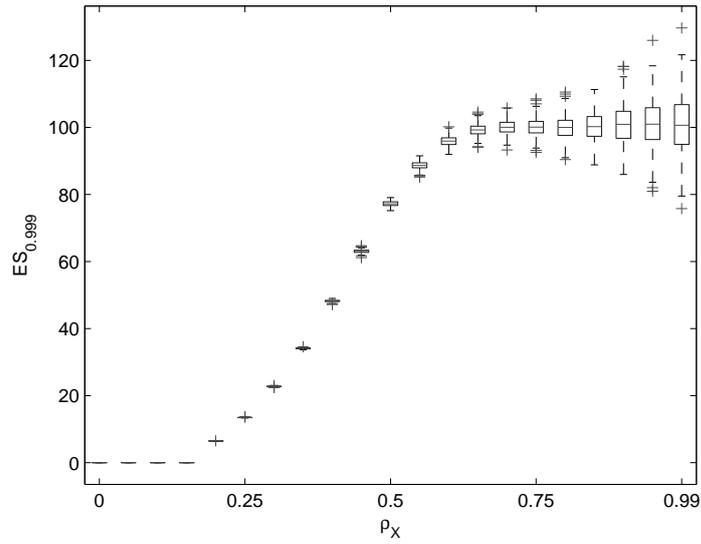
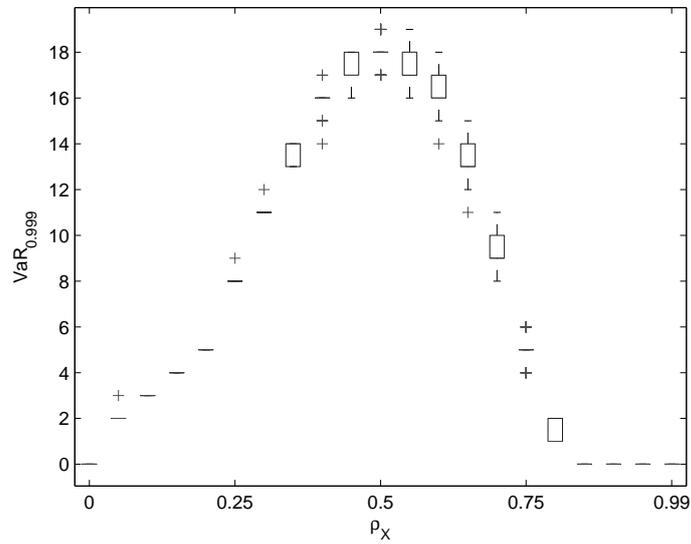


Figure 8: Mean-boxplots of 250 simulated  $\text{VaR}_{0.999}$  and  $\text{ES}_{0.999}$  figures from a Bernoulli mixture model with beta mixing distribution,  $\pi = 1.0\text{e-}005$ ,  $B$  replications

(a)  $\text{VaR}_{0.999}$ ,  $B = 1,000,000$



(b)  $\text{ES}_{0.999}$ ,  $B = 10,000,000$

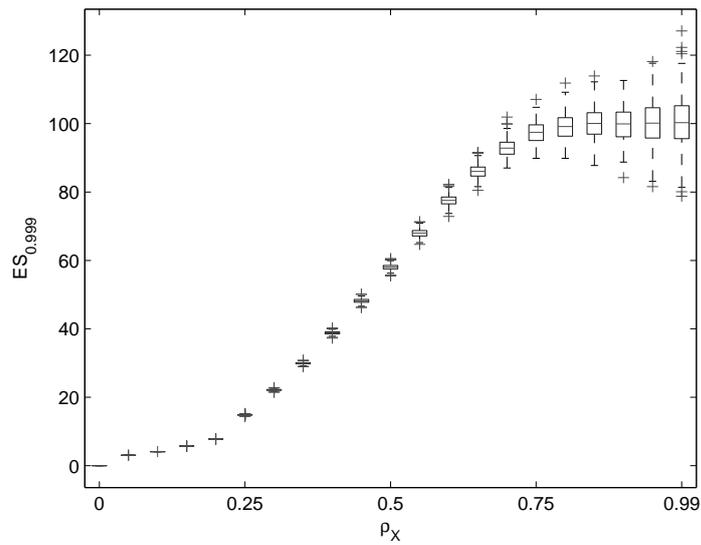
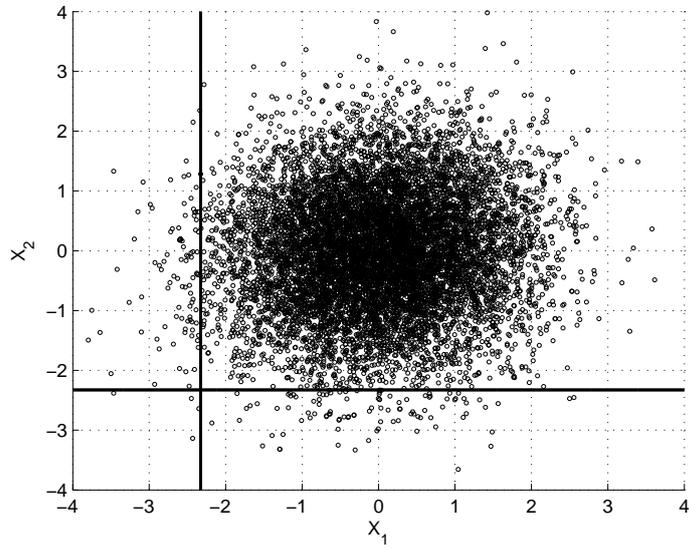


Figure 9: Mean–boxplots of 250 simulated  $\text{VaR}_{0.999}$  and  $\text{ES}_{0.999}$  figures from a Bernoulli mixture model with a Clayton copula of latent variables,  $\pi = 1.0\text{e-}005$ ,  $B$  replications

(a)  $\rho_X = 0.1$



(b)  $\rho_X = 0.9$

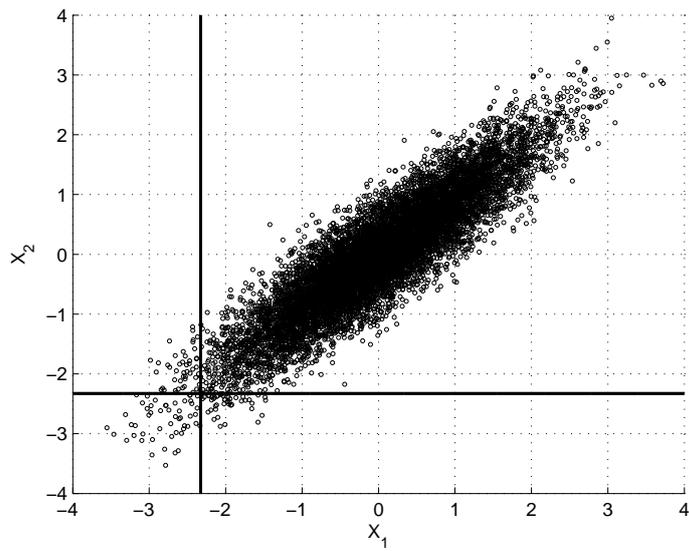


Figure 10: Scatterplots of bivariate normal distributions with  $\rho_X = 0.1$  and  $\rho_X = 0.9$