Xiang-Yi Chen

Statistical Inference of the ASRF Model in Credit Risk Management

Master Thesis, Defend on August 27, 2012

Thesis Advisor:
Prof. Dr. A. W. van der Vaart
Prof. Dr. J. J. Meulman

Specialization: Statistical Science

Mathematisch Instituut, Universiteit Leiden

Content

Summary

This paper focuses on the statistical inference of three key measures under the ASRF model framework in credit risk management: the asset correlation, PD and VaR. The first two measures are the key parameters for calculating VaR. The paper starts with some basic concept of the credit risk related to later presentation. Then, the methodology of the ASRF model and some working assumptions are elaborated in Chapter II, which also play the crucial roles in the simulation steps. Sequentially, Chapter III highlights the typical input data for estimation, and presents diverse procedures for inferring the three key measures respectively, mainly in a Bayesian framework. Later, Chapter IV demonstrates the results of estimation based on a benchmark exercise data and repeatedly simulated data sets, and comparing the performance between difference solutions. More importantly, this chapter also shows the influence of the estimation uncertainty (of ρ and PD) on measuring VaR. Finally, the conclusion is summarized in Chapter V.

Chapter I

Introduction

Risk management is at the heart of any bank's activity, since it is highly linked with the survival ability of a bank in adverse economics circles. More specifically, the managers or supervisors need to calculate how much capital should be reserved for covering the potential loss deriving from the undertaking risks. This reserved capital is called the *regulatory capital* or the *economic capital*. The former is calculated according to regulators' rules and methodologies, while the latter is an internal estimate by the bank itself. The purpose of such calculation is to use the minimum capital to cover the most of the potential loss. This optimization is necessitated due to the fact that: the more capital reserved, the more possibility to survive but less profit, and vice versa. Therefore, before a numerical analysis goes up on the stage, the manager or supervisors need to decide how much percentage of the potential loss they would like to cover. While the minimum percentage has clear reference in the authority papers (i.e. Basel II), the percentage can be set arbitrarily close to 1, according to the individual appreciation based on the sense of the risk and the expectation of the profit. For instance, if the manager wants to cover 99% percent of the potential loss, the economic capital equals to the 0.99 th quantile of the distribution of the potential loss, where the q th quantile of a random variable Y is defined as

$$\alpha_q(Y) = \inf\{l: Pr(Y \leq l) \geq q\}$$

This quantile is also widely known as *Value-at-Risk* (VaR) at 99% confidence level, since it leaves 1% worst scenario out of the consideration.

To calculate such a quantile, the essential is to quantify the distribution of the potential loss, the uncertainty of which derives from different types of the risks. Out of all, Credit risk is the single largest risk for most of the financial institutions, and it is also of our principle interests in this paper. As defined by Saita (2007), the credit risk is the risk arising from an *unexpected* deterioration in the credit quality of counterparties (i.e. borrowers/obligators). The word "unexpected" emphasizes the fact that a certain number of defaults in a portfolio are natural in the commercial context, which is expected and can be estimated beforehand. Conceptually, the potential loss of an obligator can be quantified as

(*) $$L = EAD \times PD \times LGD$$

where EAD stands for the amount of exposure (i.e. money) in the portfolio at the time default occurring, PD stands for the possibility of default and LGD stands for a percentage of the money (out of EAD) we will loss given default. The uncertainty of this quantity can be driven by both PD and LGD. This *expected loss* is usually calculated as the expectation of the potential loss, and the unexpected loss is quantified as the departure between the expected loss and a VaR at certain confidence level.

One more remark, depending on different valuation methods of EAD in (*), a credit risk management system can consider different sources of loss. There are generally two approaches for evaluating EAD: a mark-to-market (MTM) approach and a book-value accounting (BVA) approach. In the former approach, EAD need to be calculated according to its market value so that downgrades (of internal or external rating) will be considered as a source of credit loss. In the latter approach, the losses occur only when there is a reduction in the book value driven by the default of the borrowers. In this paper, we assume all the exposures are evaluated in the BVA approach, so that we can focus on the default risk.

I-1. Risk-factor Modeling

In order to quantify the potential loss caused by defualts, we first need to know what default is. For researches, a straightforward definition of default[1] is that the (logarithmic) asset value/return falls below certain threshold. At the kernel of such default mechanism, it is about the construction of the dependence structure among the asset values of obligators in a portfolio. To consider a possible source of the asset dependency, one may first think of the influence of the macroeconomic/market condition. Nowadays, all the financial individuals practice the business in a common platform. Therefore, they also share the shock of the general conditions (i.e. systematic risk), the extent of which can be variant in accord with industries and individual exposure to the platform. This can be expected to be the main source of the dependencies. Alternatively, the dependence of the defaults can also be raised by direct or indirect collaboration among obligators[2]. However, the latter source of dependence is often lack of proper measures and has some overlapping with the former source. This overlapping grows bigger if we stratify the common platform in more details. Therefore, it is then usually assumed to be negligible while the model has already considered the general economic conditions.

Typically, the dependencies among obligators in a credit portfolio model is modeled by the risk-factor framework, involving a set of latent common factors $\boldsymbol{M} = (M_1, M_2, \dots, M_p)^T$ dri-

[1] In practice, there are diverse definitions of “default” according to different regulation authorities (see Saita 2007), the financial institutions themselves can also have specific definition of default for internal exercises.
[2] The term of “Obligator” in this paper has a broad reference, such as companies, institutes, banks and any possible unit of financial individual. It can also refer to a financial instrument.

ven by diverse aspects of macroeconomics or market conditions and a idiosyncratic effect ε driven by individual financial performance. To note, the later, also called unsystematic or obligator-specific effect, represents the undiversfiable idiosyncratic risk in a portfolio, which is normally assumed to be eliminated by perfect diversification. Departure from the perfection in reality brings some undiversfiable risk, which contains nuisance information and thus is of no primary interests (comparable to the random noise in a linear regression model). Therefore, it is dubbed as an "effect", instead of the idiosyncratic risk, for the sake of clarity. More concretely, denote the logarithmic return of obligator *i* as V_i, and

(F1-1) $$V_i = \sum_{k=1}^{p} \rho_{ik} M_k + \tau_i \varepsilon_i = \boldsymbol{\rho}_i^T \boldsymbol{M} + \tau_i \varepsilon_i, \;\; i = 1,2, \dots, N$$

where the common factors are assumed to have a well defined *p*-dimensional joint distribution, $\boldsymbol{M} \sim F_M(\mu_M, \boldsymbol{\Sigma_p})$, and $\varepsilon_i \sim i.i.d\, N(0, \delta^2)$ which are independent of $\boldsymbol{M}$. Besides, $\boldsymbol{\rho}_i = (\rho_{i1}, \rho_{i2}, \dots, \rho_{ip})^T$, and ρ_{ik} describes the loading of obligator *i* on the common factor *k*, the true value of which is unknown but assumed to be non-stochastic, and τ_i is the contribution of the idiosyncratic effect. In such a framework, one can easily infer that $Var(V_i) = \boldsymbol{\rho}_i^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_i + \tau_i^2 \delta^2$ and

(F1-2) $$Corr(V_i, V_j) = \frac{Cov(\boldsymbol{\rho}_i^T \boldsymbol{M}, \boldsymbol{\rho}_j^T \boldsymbol{M})}{\sqrt{(\boldsymbol{\rho}_i^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_i + \tau_i^2 \delta^2)(\boldsymbol{\rho}_j^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_j + \tau_j^2 \delta^2)}} = \frac{\boldsymbol{\rho}_i^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_j}{\sqrt{(\boldsymbol{\rho}_i^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_i + \tau_i^2 \delta^2)(\boldsymbol{\rho}_j^T \boldsymbol{\Sigma_p} \boldsymbol{\rho}_j + \tau_j^2 \delta^2)}}$$

This correlation quantity can be greatly simplified in the one-common-factor case, or putting hypothetical restrictions on two pairs of parameters: (1) $\boldsymbol{\Sigma_p}$ and δ^2, (2) $\boldsymbol{\rho}_i$ and τ_i, which are left to the later context. According to the default mechanism, the default indicator of obligator *i* is just

(F1-3) $$I_i = 1_{\{V_i < threshold\}}.$$

The default indicators in a portfolio are not independent due to the correlated asset values, but the independence can be reached by conditioning on the common factors.

Now, we consider the total loss of a portfolio once the default indicators of all obligators are observed. Let A_i ($A_i > 0$) be the exposure of obligor *i* in the portfolio, which are assumed to be known, and U_i be the loss per euro exposure given default where $U_i \in [0,1]$. In plain terms, once the obligator *i* defaults, the investor looses $U_i A_i$. Conceptually, U_i is commonly known as the percentage *loss given default* (LGD). Additionally, the conditional independence assumption of the default events in the tracked timeline is extended to the conditional independence of U_i. Therefore, for a portfolio of *n* obligors, the portfolio loss ratio L_n is defined as

(F1-4) $$L_N = \frac{\sum_{i=1}^{N} U_i A_i \cdot I_i}{\sum_{i=1}^{N} A_i} = \sum_{i=1}^{N} w_i \cdot U_i \cdot I_i$$

such that we are free of concerns on the original exposure size and currency. The distribution of this loss quantity is mainly based on the uncertainty around the observed I_i, and also U_i can be of a stochastic property depending on the working interests. Now, we are interested in

$VaR_q(L_N)$, which is defined as the q th quantile of the distribution of L_N (i.e. $VaR_q(L_N) = \alpha_q(L_N)$). Under the model (F1-1&4) and two assumptions, Gordy (2003) proved that, as the portfolio size increases (i.e. $N \to \infty$), the q th quantile of L_N is equivalent to the q th quantile of a random quantity $E[L_N|\boldsymbol{\vartheta}]$, which is much easier to calculate.

$$VaR_q(L_N) \to VaR_q(E[L_N|\boldsymbol{\vartheta}]) \quad as \quad N \to \infty$$

$\boldsymbol{\vartheta}$ denotes a set of variables, which normally consists only $\boldsymbol{M}$ but not necessarily. The assumptions warranted for assuring such proposition are the following:

Condition (1): the $\{U_i\}$ are bounded in the unit interval and, conditional on a set of variables $\boldsymbol{\vartheta}$, are mutually independent.

Condition (2): the A_i are a sequence of positiove constants such that (a) $\sum_{i=1}^{N} A_i \uparrow \infty$ and (b) there exists a $\xi > 0$ such that $A_N/\sum_{i=1}^{N} A_i = O(n^{-(1/2+\xi)})$

The Condition (1) about U_i is in line with the definition of U_i, and the Condition (2) are sufficient to guarantee that, as the portfolio size increases (i.e. $N \to \infty$), the impact of the largest single exposure on the portfolio vanishes to zero. To note, these asymptotic properties require no restriction on the dimension of $\boldsymbol{M}$, nor any assumptions about the relationship between U_i and A_i. Therefore, they are still of great value in the multi-factor specification and the condition that, for example, the investor believes high quality loans (i.e. small U_i) tend also to be the largest loans (i.e. large A_i).For the proof of this proposition, we refer to the original paper.

Such risk-factor foundation is prevalent due to its great compatibility with current reputable industry models of portfolio credit risk, including CreditMetrics (from RiskMetrics Group), CreditRisk+ (from Credit Suisse Financial Product), CreditPortfolioView (from Mckinsey) and KMV's Porfolio Manager (Moody's/KMV). For details derivation of these industrial models on the risk-factor foundation, we refer to Gordy (2000) and Crouhy et al. (2000). As an alternative approach to the factor assumption, one can model each of the latent common factors independently by copula to generate the dependence structure, which is beyond the scope of this paper and we refer to Hardle et al. (2008) for details.

Notation

In this paper, the bold character stands for a vector or a matrix.

Chapter II

ASRF Model

On the risk-factor foundation, the simplest scenario is one common factor model, widely known as *Asymptotic Single Risk Factor* (ASRF) model, which underpins the internal-rating based (IRB) approach of Basel II. In a stylized analysis, the ASRF model commonly refers to one-factor Gaussian model with restrictive assumptions facilitating daily exercises. This is also the form how it was originally designed by Vasicek (1991). However, the framework itself is compatible to a more general context, as presented in Tarashev (2005). The "asymptotic" in the title states the ideal situation of infinite obligators in the portfolio, so that all obligators are of equal influence on the Profit/Loss of the portfolio. This property is also called *perfect fine-grained*

In this subsection, we first point out the features of the usual data for the credit risk modeling and the stylized assumptions for practical implementation. Then, we outline the methodological content of the Yearly Stochastic ASRF (YS-ASRF) model in line with the stylized features. Finally, we briefly discuss about its merits and demerits. The YS-ASRF model here is based on the model presented inTarashev (2010) with a refinement on the default modeling.

II-1. The Stylized Features and Assumptions

II-1-1. Features of the Stylized Data for Credit Risk Modeling

There are two features of the stylized data for usual modeling.

Feature 1. *The investors may observe weekly or monthly data of the asset values, but the default indicators are observed only once in a year, typically at the end of the year.*

Feature 2. *A batch of obligators is normally followed for only one year*

The first feature tries to convert such information: the observation frequency of the asset returns and the default indicators are different. The second feature depicts that the tracked batch of obligators varies every year in the timeline[3]. Therefore, the subscript *i* indexing for obliga-

[3] It is naturally but not necessary to start the observation at the beginning of calendar year. However, by assuming this, it will streamline the discussion without losing generality.

tors is independent with t which indexes the tracked years. As our target task is to analyze such stylized data sets, we expect a sounded risk-factor model should be compatible to these features. Typically, the asset correlation is inferred based on the data sets of asset returns, while PD is estimated given the asset correlation and the data sets of default indicators.

II-1-2. The Stylized Assumptions

In a stylized analysis, there are some assumptions commonly kept to facilitate daily exercise. More specifically,

(1). Although the obligators followed this year are not exactly the same people of interests in the next year, the number of obligators followed in a year stays as a constant N across the timeline. (i.e. the portfolio size N is assumed to be time invariant);

(2). All the N exposures[4] in the portfolio are of equal size, therefore, $w_i = 1/N$ for all i, which suggests fine-grained. This assumption gets closer and closer to the real situation when N increases;

(3). The common factor and the idiosyncratic effect are independent from each other, and assumed to be serially uncorrelated and both distributed as $N(0,1)$;

(4). The portfolio is homogeneous in the sense that PD and the asset correlation at a given time are the same across all the obligators. Furthermore, the PD and the asset correlation are also presumed to be time homogeneous, that is, the value of these parameters stays the same throughout the observed time horizon.

(5). LGD is assumed to be the same across the portfolio and the time. According to the regulation (Basel II), $E(U_i|M_t) = 0.45$ for all i and t, which is just a constant of no interest in the estimation procedure

II-2. Methodology of Yearly Stochastic ASRF (YS-ASRF)

The YS-ASRF first formulates the dependencies among the (log) asset values[5] with a single common risk factor, and depicts that a default event happens if the asset value of an obligator falls below a certain threshold.

Keeping the setting in line with the stylized features and assumptions, we denote the index of obligator as $i \in \{1,2,\dots,N]$ and the year index $t \in \{1,2,\dots,T]$. Furthermore, we denote that the frequency of observations in a year with equal time interval as Δ, and employ the index $h \in \{1,2,\dots,\Delta\}$. As the benchmark exercise in this paper, we assume that the investors observe the asset return on the monthly base, and therefore $\Delta = 12, h \in \{1,2,\dots,12\}$. As stated before, the obligator index i is independent with the yearly time index t, but not independent with the monthly time index h.

[4] "Obligator" and "exposure" are used as synonyms in the thesis.

[5] In this paper, we set the asset value/return in the logarithmic form everywhere.

II-2-1. YS-ASRF within certain year

In a tracked year ***t***, the YS-ASRF model builds the asset values in the manner of a stochastic process. Denote the monthly asset values of obligator *i* as $V_{i,h}$ and the jump (i.e. asset return) between two successive asset values of obligator *i* as

$$X_{i,h} = V_{i,h} - V_{i,h-1}$$

and then

(F2-1) $$V_{i,h} = V_{i,0} + \sum_{p=1}^{h} X_{i,p}$$

where $V_{i,0}$ is the initial values which are just some given constants. Assume that the behavior of $\{V_{i,h}\}_h$ is subject to a Brownian motion. More specifically, for any given *i*¸

$$X_{i,h} \sim i.i.d\ N(\mu/\Delta\,, \sigma^2/\Delta)$$

for some real μ and σ, which implies that $Corr(X_{i,h}, X_{i,h'}) = 0\ (h \neq h')$. Then, we formulate the asset return of obligator *i* at a given month *h* as

(F2-2) $$X_{i,h} = \mu \cdot \frac{1}{\Delta} + \sigma \cdot \sqrt{\frac{1}{\Delta}} \cdot \left(\sqrt{\rho} \cdot Y_h + \sqrt{1-\rho} \cdot \epsilon_{i,h}\right),\ \Delta = 12$$

where

$\sqrt{\rho}$ is the loading of obligators on the common factor, $\rho \in [0,1]$;

Y_h and $\epsilon_{i,h}$ are the monthly common factor and the monthly idiosyncratic effect;

Y_h and $\epsilon_{i,h}$ are independent and both serially uncorrelated;

$Y_h \sim i.i.d.N(0,1)$, $\epsilon_{i,h} \sim i.i.d.N(0,1)$;

$Cov(\epsilon_{i,h}\epsilon_{j,h}) = Cov(\epsilon_{i,h}, \epsilon_{i,h'}) = 0\ (i \neq j, h \neq h')$,

such that $Corr(X_{i,h}, X_{j,h}) = \rho, Corr(X_{i,h}, X_{j,h'}) = 0\ (, i \neq j, h \neq h')$. Based on (F2-1) and (F2-2), we have

(F2-3) $$V_{i,h} = V_{i,0} + \frac{h}{12} \cdot \mu + \sqrt{\frac{\sigma^2}{12}} \cdot \sqrt{\rho} \cdot \sum_{p=1}^{h} Y_p + \sqrt{\frac{\sigma^2}{12}} \cdot \sqrt{1-\rho} \cdot \sum_{p=1}^{h} \epsilon_{i,p}$$

and then, in a given month h, the marginal distribution of the asset value of obligator *i* is

$$V_{i,h} \sim N\left(V_{i,0} + \frac{h\mu}{12},\ \frac{h\sigma^2}{12}\right).$$

In a more organized form, we can model on the standardized monthly asset value $R_{i,h}$, where

(F2-4) $$R_{i,h} = \frac{V_{i,h} - V_{i,0} - {}^{h\mu}/_{12}}{\sigma\sqrt{{}^{h}/_{12}}} = \frac{\sum_{p=1}^{h} X_{i,p}\ - {}^{h\mu}/_{12}}{\sigma\sqrt{{}^{h}/_{12}}}, \text{so that marginally } R_{i,h} \sim N(0,1),$$

This also implies that, in certain month h, for any two obligator *i* and *j* in the portfolio,

$$Corr(R_{i,h}, R_{j,h}) = \frac{12}{h\sigma^2}\ Cov(\textstyle\sum_{p=1}^{h} X_{i,p}\ ,\ \sum_{p=1}^{h} X_{j,p})$$

$$= \frac{12}{h\sigma^2} \sum_{p=1}^{h} Cov(\ X_{i,p},\ X_{j,p}) = \frac{12}{h\sigma^2} \sum_h \frac{\rho\sigma^2}{12} = \rho$$

Give the above result ant that R_{i,h_t} is marginally standardized normally distributed in every month, the joint distribution of the standardized asset values in any month is then

$$\boldsymbol{R}_h = \begin{pmatrix} R_{1,h} \\ R_{2,h} \\ \vdots \\ R_{N,h} \end{pmatrix} \sim N\left(\boldsymbol{0}\ , \begin{pmatrix} 1 & & & \rho & \\ & 1 & & & \\ & & \ddots & & \\ & \rho & & 1 & \\ & & & & 1 \end{pmatrix}_{N\times N} \right) = N(\boldsymbol{0}, \boldsymbol{\Sigma}_y)$$

Finally, at the end of the year (i.e.$h = 12$), the investors will observe the dichotomous default indicators of all obligators in the portfolio, denoting the default event as 1. Recalling the default mechanism, the PD is then

$$\Pr\left(R_{i,12} < C\right) = PD$$

where C denotes the default threshold $\Phi^{-1}(PD)$.

II-2-2. YS-ASRF over T years

In order to show the formulation of YS-ASRF over the tracked *T* years, we need to use an interim time index. Denote the *h* th month in the year *t* as h_t.

Generally, the YS-ASRF model is built identically and independently for all the T years. Putting this in a more concrete way, the first target variable in the YS-ASRF model is the standardized monthly asset values (R_{i,h_t}). Based on (F2-3) and (F2-4),

(F2-5) $$R_{i,h_t} = \sqrt{\frac{1}{h_t}} \cdot \left(\sqrt{\rho} \cdot \sum_{p=1}^{h_t} Y_p + \sqrt{1-\rho} \cdot \sum_{p=1}^{h_t} \epsilon_{i,p}\right)$$

especially, denote the $R_{i,12_t}$as $R^*_{i,t}$, where

$$R^*_{i,t} = \sqrt{\rho} \cdot \left(\sqrt{\frac{1}{12}} \cdot \sum_{p=1}^{12} Y_p\right) + \sqrt{1-\rho} \cdot \left(\sqrt{\frac{1}{12}} \cdot \sum_{p=1}^{12} \epsilon_{i,p}\right)$$

(F2-6) $$= \sqrt{\rho} \cdot M_t + \sqrt{1-\rho} \cdot \varepsilon_{i,t}$$

where M_t and $\varepsilon_{i,t}$ denotes the yearly common factor and the yearly idiosyncratic effect. Assume that the serially uncorrelated property of Y_{h_t} and ϵ_{i,h_t} holds for over t years, one can see the following conditions hold:

$$M_t \sim N(0,1), \varepsilon_{i,t} \sim N(0,1), Cov\left(M_t, \varepsilon_{i,t}\right) = 0, Cov\left(\varepsilon_{i,t}, \varepsilon_{j,t}\right) = 0, \text{ for all } j \neq i \text{ and } t;$$
$$Cov(M_t, M_{t'}) = 0, Cov\left(\varepsilon_{i,t}, \varepsilon_{i,t'}\right) = Cov\left(\varepsilon_{i,t}, \varepsilon_{j,t'}\right) = 0, \text{ for all } j \neq i \text{ and } t \neq t';$$

which indicates that $Corr\left(R^*_{i,t}, R^*_{j,t}\right) = \rho,\ Cov\left(R^*_{i,t}, R^*_{i,t'}\right) = 0, i \neq j, t \neq t'$. The default indicator of obligator *i* in year *t* can then be modeled as

(F2-7) $$I_{i,t} = \begin{cases} 1, & \sqrt{\rho} \cdot M_t + \sqrt{1-\rho} \cdot \varepsilon_{i,t} < \Phi^{-1}(PD) \\ 0, & otherwise \end{cases}$$

To note, based on (F2-6), the yearly common factor is calculated as a weighted sum of the monthly common factors, and similar for the yearly idiosyncratic effect. Sequentially, (F2-7) shows that this YS-ASRF model has considered the monthly fluctuation due to the general market and individual performance when modeling the default at the end of the year.

Due to the correlated asset values, the default indicators ($I_{i,t}$) are not independent at certain *t*. However, conditional on the common factor, the default events (i.e.$I_{i,t}|M_t$) become independent in year *t*. Thus, we have

$$I_{i,t}|M_t \sim i.i.d\ Bernoulli\ (1, PD)$$

$$\begin{aligned} E(I_{i,t}|M_t) &= Pr(R^*_{i,t} < \Phi^{-1}(PD)|M_t) \\ &= Pr(\sqrt{\rho}\cdot M_t + \sqrt{1-\rho}\cdot \varepsilon_{i,t} < \Phi^{-1}(PD)|M_t) \\ &= Pr\left(\varepsilon_{i,t} < \frac{\Phi^{-1}(PD)-\sqrt{\rho}\cdot M_t}{\sqrt{1-\rho}}\middle| M_t\right) \\ &= \Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\cdot M_t}{\sqrt{1-\rho}}\right) = PD(M_t) \end{aligned} \qquad \textit{(F2-8)}$$

where $PD(M_t)$ denotes the conditional PD and it is not time-invariant. Furthermore, based also on (F1-3), at a given time *t*, the expectation of the conditional loss is

$$E(L_N|M_t) = \sum_i^N w_i \cdot E(U_i|M_t)\cdot E(I_i|M_t)$$

which has an underlying assumption that LGD (i.e.U_i) is independent of both the common factor and the idiosyncratic effect[6]. However, we will leave $E(U_i|M_t)$ out of the formulation in the later context, due to the assumption (5) in ***II-1-2*** for simplicity. Together with assumption (1), we have

$$\textit{(F2-9)} \qquad E(L_N|M_t) = \frac{1}{N}\sum_i^N \Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\cdot M_t}{\sqrt{1-\rho}}\right) = \Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\cdot M_t}{\sqrt{1-\rho}}\right)$$

which is a random variable. By the proposition of Gordy (2003) given in Chapter I, the q th quantile of the unconditional loss (L_N) is asymptotically identical to the q th quantile of $E(L_N|M_t)$. More practically, we may take a short cut by using formula (F2-6), which is also proposed by Gordy (2003) and briefly proved in **Appendix** A.

$$\textit{(F2-10)} \qquad VaR_q[E(L_N|M_t)] = E\left(L_N\middle|M_t = \alpha_{1-q}(M_t)\right) = \Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\rho}\cdot\alpha_{1-q}(M_t)}{\sqrt{1-\rho}}\right)$$

It is important to notice that (F2-8~10) are conditional only on the common factor due to the assumption that the true values of PD and the asset correlation are known. However, in practice, these two key parameters need to be estimated from the data, and therefore with uncertainty. Taking the estimation uncertainty into account, the conditional default indicator and the conditional loss should be rewritten as $I_{i,t}|M_t, PD, \rho$ and $L_N|M_t, PD, \rho$ respectively. The right sides of (F2-8) and (F2-9) stay the same under such change, but (F2-10) holds only when

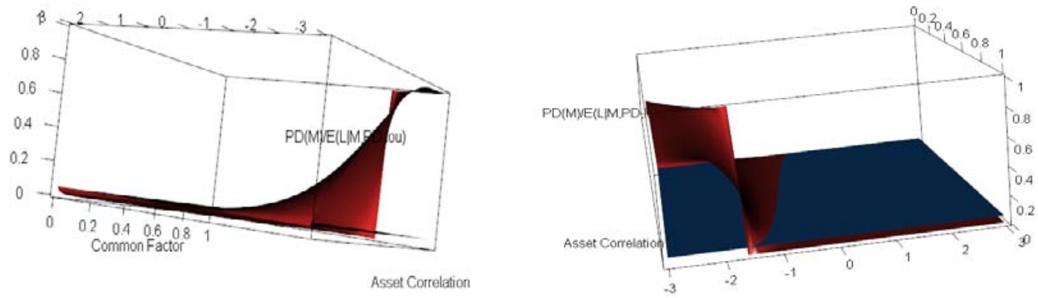
[6] Violation of this assumption is discussed in Kupiec (2008).

the estimation uncertainty is ignored. Hence, the short-cut approach for calculating VaR may bring in remarkable bias and called *naive VaR* in Tarashev (2010).

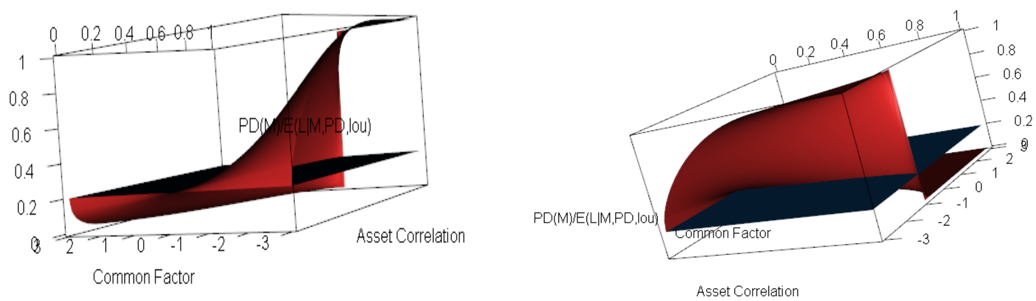
The following graphs can help to gain more insights on (F2-8) and (F2-9) when M_t, PD, ρ are all set to be viable. The axis in the horizontal panel include the common factor ($M_t \in [-3,3]$) and the asset correlation ($\rho \in [0,1]$) respectively, while the vertical axis can stand for either the conditional PD or $E(L_N|M_t, PD, \rho)$. The upper row of the graphs presents (F2-9) in 3D when setting unconditional PD = 0.05, while the bottom row of the graphs are with unconditional PD = 0.2. The blue horizontal panel represents the corresponding the unconditional PD.

When $M_t < 0$, $PD(M_t)$ and $E(L_N|M_t, PD, \rho)$ increase, especially rapidly if ρ is large. In other words, the more bad the common economics is, the more sensitive $PD(M_t)$ and $E(L_N|M_t, PD, \rho)$ are to the change of ρ. This trend becomes more obvious when the unconditional PD increases. Besides, the graphs shows that, when $-1 < M_t < 0$ but ρ is non-zero, the conditional PD is smaller than the unconditional PD. This indicates that the strong asset correlation is actually helpful under a mild challenging situation. However, the high asset dependency becomes harmful when the general economic keeps falling down. This is due to the fact that, once an obligator suffers bankruptcy, her/his close business partners get hinged hits. One the other hands, when $M_t < 0$, $PD(M_t)$ and $E(L_N|M_t, PD, \rho)$ are generally small. Hence, the asset correlation has no significant impact in this situation.

Unconditional PD=0.05



Unconditional PD=0.20



II-3. Merits and Demerits of ASRF

Despite elegant mathematical aspects, the reason for choosing ASRF is two-fold: 1) the ASRF model allows the investors and manages to calculate VaR based on their internal rating and information (IRB approach in Basel Committee on Banking Supervision, 2005), and 2) the ASRF framework eases the simulation procedure and computational burden for estimation, and hence efficient for the daily practice.

However, one might find the ASRF framework is very restrictive with respect to the assumptions. While some working assumptions, such as homogenous ρ, can be loosen according to individual interests, the essential assumptions (i.e. single risk factor and perfect granularity) are not touchable. In a paper of Tarashev and Zhu (2008), they classified the errors of measuring the credit risk by ASRF into two categories: specification error and calibration error.

The specification error refers to the influence on the capital measures when one of the two key assumptions of this model is violated. The violation of perfect granularity was found in general having a negative impact, and this impact decreased when the number of exposures increases. For making up this shortage, Gordy et al. (2007) has derived an adjustment. For the violation of single risk factor assumption, it was found also having a negative effect, since the single risk factor assumption leads to an underestimation of the desired capital when there are multiple clusters of defaulting sources. However, in general, the specification error was concluded as being virtually inconsequential by Tarashev et al (2008).

The calibration error refers to the influence on the capital measures when the estimation uncertainty of key parameters (i.e. the asset correlation and PD) is ignored. This is the focus in our paper, and the result is presented in Chapter IV.

Chapter III

Estimation Procedure

Under the paradigm of the stylized YS-ASRF model, the targets of the estimation procedure are three measures (i.e. the asset correlation, PD and VaR). The input for the whole procedure are two panels of data: (1) a $N \times 12T$ matrix of the asset values of N obligators over *12* months of T independent cohorts and (2) a $1 \times T$ vector of default rates of N obligators in each of T years.

Conceptually, the asset correlation is purely inferred based on the panel of the asset values, while the inference of PD can only be based on a set of default rates and the estimated asset correlation. Finally, the naïve and correct VaR are calculated based on the estimates of asset correlation and PD. The reason for not using the asset returns information when deriving PD estimate is to be compatible to the real-life application: Heitfield (2008) reported that usually asset-return and default-rate data cover different sets of obligators. Figure 1 portrays the hierarchical structure of this estimation procedure. The details of the estimation procedure for these three key measures will be articulated in the next three sections (***III-1~3***), respectively.

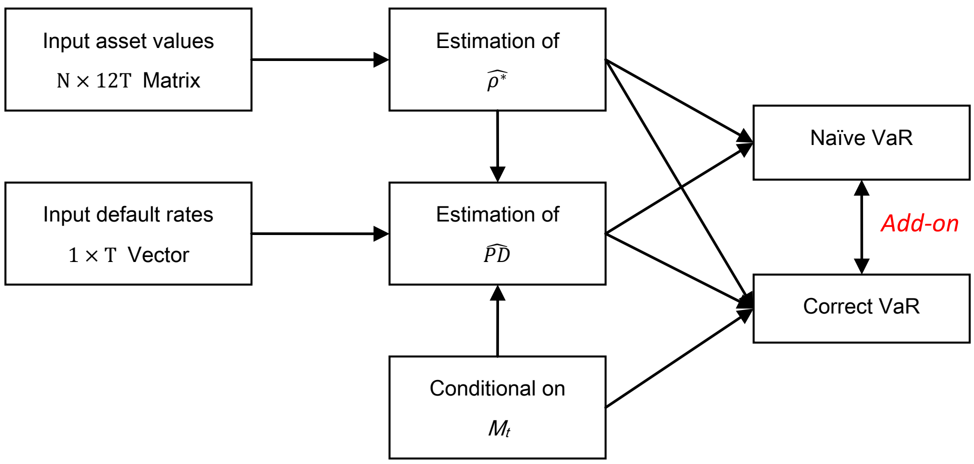


Figure 1. Graphical Expression of the whole estimation procedure

III-1. Estimation of Asset Correlation

Assume that the investors have sufficient knowledge about how the asset values are generated. In other words, they are aware of the YS-ASRF model. Therefore, the investor may model on the jump matrix instead on the asset value matrix. Denote the jump matrix over T year as $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_\mathrm{T})$ of dimension $N \times 12T$, where $\mathbf{X}_\mathrm{t}$ is a yearly jump matrix of dimension $N \times T$. For the simplicity, denote a new index $k \in \{1,2,\ldots,12T\}$. There are two solutions to the estimation of the asset correlation: a hierarchical model and a linear mixed model (LMM).

III-1-1. Tarashev Solution

Tarashev (2010) proposed the following Bayesian estimation procedure in a hierarchical manner. First of all, we need to derive a point estimate of the asset correlation ($\hat{\rho}$) based on **X**. One may simply calculate it in the following way.

(F3-1) $$\hat{\rho} = \frac{\mathbf{1}^T \widehat{\boldsymbol{\Sigma}_y} \mathbf{1} - tr(\widehat{\boldsymbol{\Sigma}_y})}{N(N-1)}$$

where **1** is a vector of all one (of dimension $N \times 1$), and $\widehat{\boldsymbol{\Sigma}_y}$ is the sample correlation matrix with respect to N rows of jump matrix $\mathbf{X}$ and a unbiased estimator of $\boldsymbol{\Sigma}_y$ in (F2-2). $\widehat{\boldsymbol{\Sigma}_y}$ can be calculated as the following.

$$\widehat{\boldsymbol{\Sigma}_y} = \frac{1}{N-1}\left(\mathbf{X}^*\mathbf{X}^{*T}\right) \quad where \quad \mathbf{X}^* = \begin{pmatrix} \mathrm{X}^*_{1,1} & \cdots & \mathrm{X}^*_{1,12T} \\ \vdots & \ddots & \vdots \\ \mathrm{X}^*_{N,1} & \cdots & \mathrm{X}^*_{N,12T} \end{pmatrix} \quad and$$

$$\mathrm{X}^*_{i,k} = \frac{\mathrm{X}_{i,k} - \sum_{k=1}^{12T} \mathrm{X}_{i,k}/12\,T}{\sqrt{\frac{1}{12T-1}\sum_{k=1}^{12T}\left(\mathrm{X}_{i,k} - \sum_{k=1}^{12T}\mathrm{X}_{i,k}/12\,T\right)^2}}$$

Then, the point estimate $\hat{\rho}$ is used as the data for a Bayesian procedure. Denote the correct value of $\hat{\rho}$ as ρ^*, correct in the sense that $\hat{\rho}$ fluctuates around ρ^* given different data sets. As a working model, assume that $\hat{\rho}$ is sampled from a Beta distribution with parameters $\left(\alpha(\rho^*), \beta(\rho^*)\right)$, for $\alpha(\rho^*)$ and $\beta(\rho^*)$ solving

$$\frac{\alpha(\rho^*)}{\alpha(\rho^*)+\beta(\rho^*)} = \rho^* \text{ and } \frac{\alpha(\rho^*)\beta(\rho^*)}{\left(\alpha(\rho^*)+\beta(\rho^*)\right)^2\left(\alpha(\rho^*)+\beta(\rho^*)+1\right)} = \frac{2(1-\rho^*)^2[1+(N-1)\rho^*]^2}{TN(N-1)}.$$

The left sides of these two equations are the first two moments of the Beta distribution with parameters $\left(\alpha(\rho^*), \beta(\rho^*)\right)$, while the right sides are chosen so that these moments are approximately equal to $E_{\rho^*}\hat{\rho}$ and $Var_{\rho^*}\hat{\rho}$ respectively. To note, the determination of

$$Var_{\rho^*}\hat{\rho} = \frac{2(1-\rho^*)^2[1+(N-1)\rho^*]^2}{TN(N-1)}$$

is presented in Tarashev (2010). Furthermore, since there is no prior information about the true asset correlation, we assign a non-informative priori $g(\rho^*)$ for ρ^*, say a uniform distribution $U(0,1)$.

The posteriori of the asset correlation following the above procedure can be summarized as the following:

$$prior:\ \rho^* \sim Uniform(0,1)$$

$$Samples:\ \hat{\rho}|\rho^* \sim Beta\big(\alpha(\rho^*), \beta(\rho^*)\big),$$

$$where\ \alpha(\rho^*) = \frac{\rho^{*2}[TN(N-1)]}{2(1-\rho^*)[1+(N-1)\rho^*]^2} - \rho^*\ \ and\ \ \beta(\rho^*) = \frac{1-\rho^*}{\rho} \cdot \alpha(\rho^*)$$

$$posterior:\ f(\rho^*|\hat{\rho}) = \frac{g(\rho^*)f(\hat{\rho}|\rho^*)}{\int g(\rho^*)f(\hat{\rho}|\rho^*)d\rho^*}$$

Implementation Note:

Instead of quoting the form of $Var_{\rho^*}\hat{\rho}$ in Tarashev (2010), we may try a more flexible setting. Let $Var_{\rho^*}\hat{\rho} = s^2$, where s^2 is some finite and non-zero real number. Now, the parameters of the Beta distribution can be rewritten as some functions of ρ^* and s^2, as the following.

$$\alpha' = \frac{\rho^{*2}(1-\rho^*) - \rho^* s^2}{s^2} \text{ and } \beta' = \frac{\rho^*(1-\rho^*)^2 - (1-\rho^*)s^2}{s^2}$$

The Bayesian procedure can be implemented after assigning a diffuse prior distribution for s^2. However, it is important in this case to realize that s^2 has a restrictive range due to (a) the requirements of $\alpha > 0\ and\ \beta > 0$ and (b) $\rho^* \in [0,1]$. It is obvious that the first condition will be met if and only if $\rho^*(1-\rho^*) > s^2$. The left side of this inequality is a convex quadratic function, and has a range of [0, 0.25] due to the domain of ρ^*. Hence, the prior of s^2 can be decided only after ρ^* is given. This Bayesian procedure can be summarized as:

$$prior:\ \rho^* \sim Uniform(0,1)$$

$$prior:\ s^2|\ \rho^* \sim Uniform\big(0, \rho^*(1-\rho^*)\big)$$

$$Samples:\ \hat{\rho}|\rho^*, s^2\ \sim Beta(\alpha', \beta'),$$

$$posterior:\ f(\rho^*|\hat{\rho}, s^2) \propto\ g(\rho^*)g(s^2|\ \rho^*)f(\hat{\rho}|\rho^*, s^2)$$

However, WinBUGS has computational difficulty for this setting. Alternatively, one may directly assign diffuse priori for the parameters of the Beta distribution (α, β), and then use their posteriori to derive the posterior estimate of ρ^*, based on the function that

$$f(\rho^*|\alpha, \beta) = \frac{\alpha}{\alpha + \beta}$$

Nevertheless, the prior of $\rho^*|\ \alpha, \beta$ in such case may not be flat.

III-1-2. LMM Solution

Other than putting a hypothetical distribution for $\hat{\rho}$, it is more straightforward to build the Bayesian estimation as the following

(F3-2) $$f(\rho^*|\mathbf{X}) \propto g(\rho^*)f(\mathbf{X}|\rho^*)$$

where each column of $\mathbf{X}$ can be considered as a realization of a random vector $\boldsymbol{X}_{h_t} = (X_{1,h_t}, X_{2,h_t}, \dots, X_{N,h_t})^T$ since $\{\boldsymbol{X}_{h_t}\}$ are mutually independent across the tracked years. Based on (F2-2), it is clear that

$$\boldsymbol{X}_{h_t} \sim N_N\left(\frac{\mu}{12}\cdot \mathbf{1}_{N\times 1}, \frac{\sigma^2}{12}\cdot \boldsymbol{\Sigma}_y\right)$$

The off-diagonal elements of $\boldsymbol{\Sigma}_y$ all equals to ρ, i.e $Corr\left(X_{i,h_t}, X_{j,h_t}\right) = \rho\ (i \neq j)$. This estimation can be formulated in LMM framework as the following.

(F3-3) $$X_{i,k} = u + b_k + e_{i,k}, i \in \{1,2,\dots,N\}$$

where b_k and $e_{i,k}$ are independent, $e_{i,k} \sim N(0,\sigma_e^2), Cov\left(e_{i,k}, e_{j,k}\right) = Cov\left(e_{i,k}, e_{i,k'}\right) = 0$, $b_k \sim N(0,\sigma_b^2), Cov(b_k, b_{k'}) = 0,$ for all $i \neq j, k \neq k'$. Comparing to the formulation of YS-ASRF model, we expect that $u, b_k, e_{i,k}$ reflects $\mu/12, Y_{h_t}, \epsilon_{i,h_t}$ respectively. Based on (F3-3),

$$Cov\left(X_{i,k}, X_{j,k}\right) = Cov\left(b_k + e_{i,k}, b_k + e_{j,k}\right) = Cov(b_k, b_k) = \sigma_b^2$$

$$Var\left(X_{i,k}\right) = Var\left(b_k + e_{i,k}\right) = \sigma_b^2 + \sigma_e^2$$

and then

(F3-4) $$Cor\left(X_{i,k}, X_{j,k}\right) = \frac{\sigma_b^2}{\sigma_b^2+\sigma_e^2} = \rho^*$$

Implementation Note:

The Bayesian inference for this LMM can be summarized as the following:

$$prior:\ u \sim\ Normal(0, 1000)$$
$$prior:\ \rho^* \sim\ Unif(0,1)$$
$$prior:\ \sigma_b^2 \sim\ Inverse\ Gamma(0.001, 1000)$$
$$Samples:\ b_k|\sigma_b^2 \sim N(0,\sigma_b^2)$$
$$Samples:\ X_{i,k}|u, b_k, \sigma_b^2, \rho^* \sim N(u + b_k, \frac{\sigma_b^2}{\rho^*} - \sigma_b^2)$$

However, in such formulation, WinBUGS recognizes more than 100 nodes, since the prior of σ_e^2 is calculated based on σ_b^2 and ρ^*. It runs very slowly and has a poor mixing rate. Therefore, we might consider to give a non-informative prior for σ_e^2 directly, and calculate the posterior estimate of ρ^* based on the posteriori of σ_b^2 and σ_e^2.

III-2. Estimation of PD

The input for PD estimation is a $1 \times T$ vector of default rates, given the estimate of asset correlation. Recall that N default indicators in a certain year have underlying dependency, which stems from the asset correlation among the obligators and causes the major difficulty of the estimation procedure. A natural question attached with this statement is about the relationship between the default correlation and the asset correlation, discussion on which is referred to **Appendix B** in full details. Fortunately, such doubt have zero impact in the formulation stage of the PD estimation.

A sound way to estimate the unconditional PD is first to settle the conditional PD, which is of the form as (F2-8), and then

(F3-5) $$PD = \int [PD(M_t)]\, d\Phi(M_t) \; .$$

where PD is presumed to be time invariant, but $PD(M_t)$ is not, with the conditional form to emphasize this. However, $PD(M_t)$ is still supposed to be homogeneous across the portfolio in a given year. Thus, The default rate (I_t) in a given year *t* is

$$I_t = \sum_i I_{i,t} | M_t \sim Binary\big(N, PD(M_t)\big)$$

and the likelihood function can be written as

$$Lik(PD(M_t)) = f\big(I_{1,t}\,, \dots, I_{N,t} \big| M_t\big) = \binom{N}{I_t} PD(M_t)^{I_t} [1 - PD(M_t)]^{N - I_t}$$

As proposed in McNeil (2003), Generalized Linear Mixed Model (GLMM) with probit link function is a good match for this case, and it can be extendedly used in the case of multiple observed and latent common factors. For details, we refer to the original paper.

In general, GLMM can be used for the following longitudinal scenario. There are several subjects in the study for, say, blood pressure. The blood pressure of each subject is recorded repeatedly across the tracked timeline as a categorical data (for example, categorical as abnormal/normal). It is then natural to expect that such repeat measures of each subject have some underlying dependency due to individual physical situation across the time. GLMM can be employed to model the possibility of having abnormal blood pressure with this latent dependency by *random effects* and possibly some interested observable covariates in *fixed effects* portion (i.e. age, gender and so on). Back to the case of ASRF model, apparently, there are no fixed-effect covariates, since the true values of the common factor and the idiosyncratic effect are all unobservable. More importantly, the *serial dependency* of the above toy example is not existed on the ASRF foundation. Instead, we are interested to formulate the dependency among obligators in a given year, and the cohorts of obligators in different years are assumed to be independent. Therefore, GLMM will take the recorded years as “subjects” and focus on the *cross-sectional dependency* of the N observations in each year. More concretely, the GLMM model in our case can be written as

(F3-6) $$E(I_t|b_t) = PD(M_t)$$

(F3-7) $$probit\big(PD(M_t)\big) = \Phi^{-1}\big(PD(M_t)\big) = X\mu + \theta b_t$$

where μ is a fixed-effect intercept, b_t is a random intercept and $b_t \sim i.i.d\ N(0,1)$, X and θ are known designed coefficients. Moreover, b_t is expected to represent M_t and μ is expected to represent the time invariant default threshold $\Phi^{-1}(PD)$, when X and θ are parameterized as

$$\theta = -\sqrt{\frac{\rho}{1-\rho}}\ ,\ X = \frac{1}{\sqrt{1-\rho}} = \sqrt{1+\theta^2} \text{ and } \mu = \Phi^{-1}(PD).$$

since, based on (F2-8), we have

$$\Phi^{-1}\big(PD(M_t)\big) = \frac{\Phi^{-1}(PD) - \sqrt{\rho}\cdot M_t}{\sqrt{1-\rho}}$$

$$= \frac{\Phi^{-1}(PD)}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}} M_t$$

There are few remarks warranted for the usage of such estimation procedure:

First of all, under such parameterization, an estimate of PD is given by the estimate of $\Phi(\mu)$ in (F3-7), dubbed as *Formal Estimate of PD*. The uncertainty around μ can also be easily transferred $\Phi(\mu)$ due to the monotonic property of the function. In a more structured way, this target estimate can be computed simultaneously with other parameters in the GLMM model via the MCMC method. An alternative way to derive an estimate of PD is averaging over the posterior distributions of the conditional PD, based on (F3-5), dubbed as *Alternative Estimate of PD*.

Secondly, X and θ are known, in the sense that we assume ρ is given beforehand. This is hardly true. In industrial practice, one may take a point estimate $\hat{\rho}$ as the true value and place it in the above model. In a more sound way, one can put a prior on ρ to incorporate with the estimation uncertainty. The posterior $f(\rho^*|\mathbf{X})$ derived in ***III-1*** can be a suitable prior in this case, so that we also use the information of asset values when inferring PD and in turn expect a more accurate estimate. In such way, the estimation procedure of the asset correlation and PD can then be integrated as a hierarchical model.

III-3. Estimation of VaR

Under the ASRF framework, VaR estimate is built based on two key parameters: the asset correlation and PD, the true values of which are unknown. For practical exercises, the estimated values of these two parameters are used in the calculation of VaR. Nevertheless, ignorance of the uncertainty of the estimates will certainly bring in bias. The bias can be quantified as the difference between the correct VaR and the naïve VaR. The correct VaR takes the estimation uncertainty into account, while the naïve VaR sets the estimated values as the true values. The difference between the two is named *VaR Add-on*.

VaR Add-on = Correct VaR – Naïve VaR

As discussed shortly in ***II-2-2***, taking the estimates of PD and the asset correlation as the true values, $VaR_q(L_N) \to VaR_q\big(E(L_N|M_t)\big)$ while $N \to \infty$, and the naïve VaR can be calculated based on (F2-10). Since $M_t \sim N(0,1)$, the naïve VaR at *q* th confidence level is more specifically calculated as

(F3-8) $$VaR_q^{naive} = \Phi\left(\frac{\Phi^{-1}(\widehat{PD}) - \sqrt{\hat{\rho}} \cdot \Phi^{-1}(1-q)}{\sqrt{1-\hat{\rho}}}\right).$$

On the other hand, the correct VaR takes account of parameter uncertainty by treating that (*PD*, *ρ*) also as random variables, so that $VaR_q(L_N) \to VaR_q\big(E(L_N|M_t, PD, \rho)\big)$. In other words, this implies that the investors are facing with multiple *common risk factors*: the common credit-risk factor (M_t) and the common estimation-risk factors (*PD*, *ρ*). Moreover, for a working model (Tarashev 2010),

Assum. 1. the joint distribution of (PD, ρ) is assumed to be well defined;

Assum. 2. Since the uncertainty of M_t *comes from the future situation while the uncertainty of (PD, ρ) is driven by the past information, the serial independent assumption of* M_t *can be extendedly narrated as the independence of* M_t *and (PD, ρ).*

The situation of multiple risk factors violates the key assumption of the ASRF framework, which invalidates the legitimacy of using (F2-10) as a short cut. The correct VaR is the *q* th quantile of the random variable $E(L_N|M_t, PD, \rho)$.

(F3-9) $$VaR_q^{correct} = \alpha_q\big(E(L_N|M_t, PD, \rho)\big),\ where\ E(L_N|M_t, PD, \rho) = \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot M_t}{\sqrt{1-\rho}}\right)$$

This quantity is of no analytical expression but can be derived via a simulation-based procedure presented as the following.

Step-1. *Determine the joint distribution of* (PD, ρ)*;*

Step-2. *Simulate a pair of* $(PD, \rho)^{[i]}$ *based on the joint distribution, and draw a large amount numbers (in our case 1000) of* M_t *from* $N(0,1)$*;*

Step-3. *Calculate the corresponding* $E(L_N|M_t, PD, \rho)$ *given 1000 simulated values of* M_t *and a pair of* $(PD, \rho)^{[i]}$*;*

Step-4. *Repeat Step-2 and Step-3 a large number times (in our case 1000, hence,* $i \in \{1,2,\dots,1000\}$*), and aggregate the results (*1000×1000 *values) to form the distribution of* $E(L_N|M_t, PD, \rho)$*. The* $VaR_q^{correct}$ *is the q th quantile of this distribution.*

Chapter Ⅳ

Calibration Errors in ASRF-based Measures

The simulated data sets are used for the estimation procedure discussed in the previous chapter, so that we can better evaluate the accuracy of estimates. The simulation scheme based on the YS-ASRF model is presented in details in ***IV-1***, where readers can expect a clear image of the data sets used in practice for estimation. Besides, we explore more interesting relationships between the unconditional PD and other setting in ASRF model (i.e. the asset correlation, portfolio size, and observed time horizon) via repeating such simulation procedure in ***IV-2***. Afterwards, we demonstrate the estimation results of the asset correlation, PD and VaR measures in ***IV-3~5***, respectively.

IV-1. Simulation Scheme

The simulation should be in line with the YS-ASRF model discussed in (***II-3***). Before going into a detailed description of simulation procedure, we first scrutinize the stylized data sets again, which sets the target for our simulation.

Under the stylized framework, the investors track the asset values of N obligators for a year, and 12 observations of each obligator are recorded monthly in the year. By the end of the year, the investors also observe the default rate, which is the sum of the default indicators in that year from the simulation perspective. The tracking procedure is repeated over T years on different batches of obligators, in order to cumulate enough information to run the inference with acceptable accuracy. In other words, the N obligators tracked in different years are different, so that they can be treated as T independent cohorts ($N \times T$ obligators in total). Accordingly, we can expect to observe a panel data of asset returns (of dimension $N \times 12T$) and default rates (of dimension $1 \times T$) for credit risk assessment. The first target data set can be viewed as a horizontal-augmenting matrix by merging T different $N \times 12$ matrixs. Moreover, three further remarks are warranted for a proper usage of such stylized data sets:

(a). The frequencies of asset value and default rate observations are in line with common financial exercises. Besides, recording asset values monthly can filter out high-frequency noise. There is one more point should be kept in the reminder. The asset values seem to be easily accessed so far, while they are not directly observable in real world. Usually, the asset values are extracted from selected market data, such as stock price and so on, and confounded with noise thereby.

(b). Despite the theoretical realization of asymptotic properties of ASRF requires infinite obligators in the portfolio, N is always finite in reality. For a benchmark exercise suggested in Tarashev (2010), N is set as 200. In this paper, for any sensitivity analysis related to N, we will consider a candidate-value pool of {50, 100, 200, 500, 1000}. In addition, N is suggested as time invariant in the paper for the illustrative purpose.

(c). Basel II accord regulated banks should base their PD estimates on observations over at least 5 years. In the paper, for any sensitivity analysis related to T, we will consider a candidate-value pool of {5, 10, 20, 35, 50 (unit in years)}. It is barely possible for existing financial institutes to have unified records over 50 years, but it is helpful for demonstration.

(d). From the simulation point of view, we can observe the default indicators of each obligators in a given year. However, we only take the default rate (the sum of these default indicators) as the input for the PD estimation, in order to emulate the common practices (described in Paragraph 2, Chapter III).

Taking the above voluminous profile of the stylized data along the side, we now construct a simulation scheme leading to the two target data sets. First of all, we simulate the initial values for each obligators in certain year t, and then simulate the $N \times 12$ matrix of asset values based on (F2-1), the essential of which falls on independently 12 times simulation of $\boldsymbol{X}_{h_t}$. Based on (F2-2), the simulation of $\boldsymbol{X}_{h_t}$ can be decomposed into two steps: 1) simulate the monthly common factor Y_{h_t} and the monthly idiosyncratic effect ϵ_{i,h_t} independently, and then 2) take the weighted sum, as the following. For each month ($h_t \in \{1,2,\dots,12\}$) in year t,

$$\boldsymbol{F}_{(N+1)\times 1} = \begin{pmatrix} Y_{h_t} \\ \epsilon_{1,h_t} \\ \vdots \\ \epsilon_{N,h_t} \end{pmatrix} \sim N_{N+1}\left(\boldsymbol{0}, \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}_{N+1} \right)$$

and then $\boldsymbol{X}_{h_t} = \boldsymbol{A}_{N\times(N+1)}\boldsymbol{F}_{(N+1)\times 1} + \frac{\mu}{\Delta}\cdot \mathbf{1}_{N\times 1}$, where

$$\boldsymbol{A}_{N\times(N+1)} = \begin{pmatrix} \sqrt{\rho} & \sqrt{1-\rho} & & & \\ \sqrt{\rho} & & \sqrt{1-\rho} & & \\ \vdots & & & \ddots & \\ \sqrt{\rho} & & & & \sqrt{1-\rho} \end{pmatrix}.$$

The default indicator matrix (of dimension $N \times T$) can be calculated as (F2-6) and (F2-7), and the default rates are the column sums of the matrix. Detailed simulation procedure is listed in Table 1.

Table 1. Simulation Scheme (1)

Step	Instruction
1	Define portfolio size *N*, observed time horizon *T*, and the true value of ρ and *PD*;
2	Assign proper values for μ and σ, simply by professional knowledge in financial industry, or by trial-and-error procedure until the data set of asset values over time shows a satisfactory shape (i.e. averages and the fluctuation);
3	Simulate initial asset values for *N* obligators ($V_{i,0}$) from a proper distribution at a given year. An appropriate distribution should generate *N* values with mild dispersion in order to keep in line with real situations. We choose N(23,1) for all *T* years;
4	Simulate a random number for Y_{h_t} and *N* random number for ϵ_{i,h_t}, both from standard normal distribution. Calculate the jump value $X_{i,h_t}$and the current asset value V_{i,h_t} based on (F2-1&2);
5	Repeat Step-4 12 times;
6	Calculate $R_{i,12_t}$ accord to (F2-6), compare the value with $\Phi^{-1}(PD)$, and then calculate $I_{i,t}$ as (F2-7). The default rate in year *t* is then $I_t = \sum_i I_{i,t}$;
7	Repeat Step-3~5 *T* times

Benchmark Exercise Data

In this paper, the *Benchmark Setting* is

$$N=100,\ T=10,\ \rho=0.25,\ PD=0.1,\ \mu=6,\ \sigma=15.$$

The simulation scheme is executed in R with the random seed "20120827". The simulation produces a panel of 100×120 asset returns and 10 default rates (out of 100). More specifically, the 10 default rates are listed in Table 5, together with latent yearly common factors. The average M_t is around -0.32 and therefore the general economics in these virtual 10 years is slight negative. One can see that in a good commercial timing (i.e. $M_t > 0$), say year 2, 6 and 7, the default rates are generally small. One might try to set the random seed to "20120723" and will find out, if M_t is larger than 1, there is a great chance to observe zero default.

Table 2. Benchmark Simulation

Year	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
I_t	17	6	8	20	7	4	1	10	12	9
M_t	-0.85	0.15	-0.113	-1.51	-0.42	0.27	0.98	-0.31	-0.90	-0.48

IV-2. A Simulation-based Exercise

Before getting into the estimation results following the procedure discussed in Chapter III, we can use a simulation-based exercise to explore some interesting relationship between a crude estimate of the unconditional PD and other settings in Step-1 of the simulation scheme, under the ASRF framework. This exploration will also facilitate a latter exercise (in ***IV-5***), which illustrates how the uncertainty of the asset correlation and PD influences VaR measures.

For the moment, we assume that the investors have neither any prior information about the asset correlation nor any observations of the asset values during the tracked years. Therefore, we only have default rates (I_t, defined in ***III-2***) to estimate the unconditional PD. Intuitively, a point estimator of PD is

$$Empirical\ PD = \sum_t I_t / NT \quad .$$

Given the setting in Step-1, we can obtain the distribution of such empirical PD by repeating the simulation a large number of times, say 1000 times. Afterwards, this procedure is carried on under different settings of Step-1. The results of this exercise are presented graphically in Figure 3~5 and numerically in Table 3~5.

Generally, the mean of the empirical PD is very close to the truth value of the unconditional PD. Holding other settings, the variance of such estimate decreases when *T* grows larger or the asset correlation becomes smaller, and so do the skewness and the kurtosis. In other words, we may expect the estimate of PD to be more accurate and stable when the tracked time horizon is long enough or the asset correlation is low. Moreover, the skewness of the empirical PD reduces rapidly when increasing *T* up to 20, and slowly afterwards. As the true value of PD increase, 1) the variance of the empirical PD averagely increases while the skewness decreases; 2) the change of the skewness becomes more sensitive to the change of the asset correlation. Take an instance for the latter comment, the average increases of the skewness when the asset correlation increases by 0.1 is larger in the occasion of a smaller true value of PD. Besides, one can see that, when the true value of PD equals to 0.05, the skewness increases by 0.303 if the asset correlation changes from 0.1 to 0.2. On the other hand, when the true value of PD equals to 0.1, the change of similar scale of the skewness will only happen when changes from 0.2 to 0.3.

Throughout the exercise, there is no evidence that supports the portfolio size has influence on the distribution of the empirical PD. This phenomenon is caused by the dependency among the default within a year. If we set the true value of the asset correlation to zero and run the procedure again, it returns with a boxplot, in which the variance of the empirical PD estimate decreases as the portfolio size increases. This finding is also consistent with Dwyer (2007).

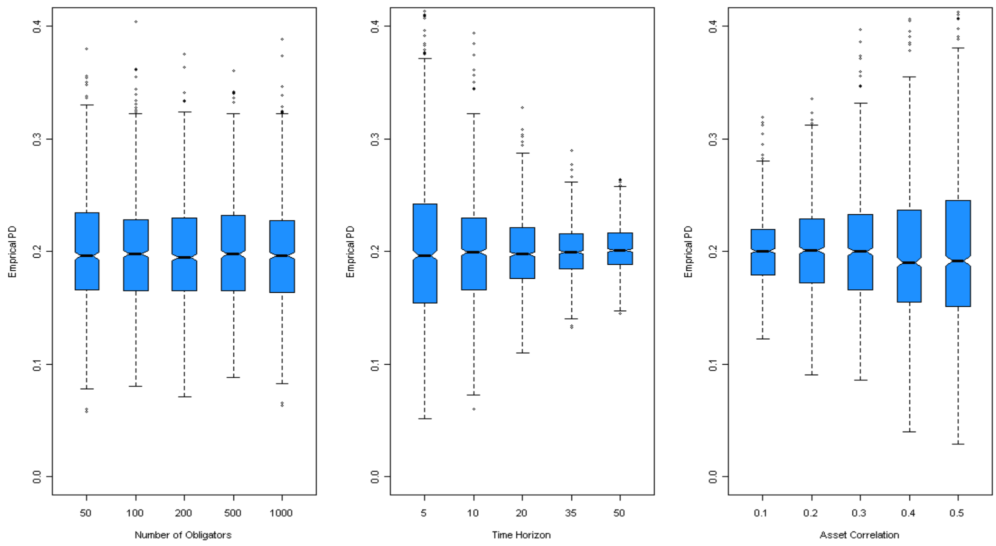


Figure 3. The Empirical PD vs. Portfolio Size (the left, holding other setting as the benchmark setting), Tracked time horizon (the middle, holding other setting as the benchmark setting), the Asset correlation (the right, holding other setting as the benchmark setting). The true value of PD = 0.20; Empirical PD are calculated based on 1000 repeated simulated data sets,

Table 3. Empirical PD (True Value of PD = 0.20)

		Mean	**Median**	**SD**	**Skewness**	**Kurtosis**	**[P95 , P5]**
	50	0.200	0.196	0.049	0.332	3.102	[0.126 , 0.284]
	100	0.199	0.198	0.048	0.391	3.274	[0.127 , 0.281]
N	*200*	0.198	0.194	0.047	0.331	3.108	[0.125 , 0.282]
	500	0.200	0.198	0.048	0.308	2.754	[0.126 , 0.282]
	1000	0.198	0.197	0.048	0.32	3.159	[0.124 , 0.282]
	5	0.202	0.196	0.065	0.515	3.359	[0.106 , 0.316]
	10	0.201	0.199	0.047	0.416	3.456	[0.131 , 0.283]
T	*20*	0.199	0.197	0.033	0.294	3.064	[0.147 , 0.256]
	35	0.200	0.199	0.024	0.165	3.081	[0.160 , 0.240]
	50	0.203	0.201	0.021	0.208	2.818	[0.169 , 0.239]
	0.1	0.200	0.200	0.030	0.359	3.390	[0.155 , 0.250]
	0.2	0.203	0.201	0.042	0.324	2.778	[0.140 , 0.280]
ρ	*0.3*	0.202	0.200	0.052	0.378	3.087	[0.123 , 0.291]
	0.4	0.196	0.190	0.060	0.472	3.416	[0.104 , 0.301]
	0.5	0.200	0.192	0.069	0.461	3.078	[0.099 , 0.321]

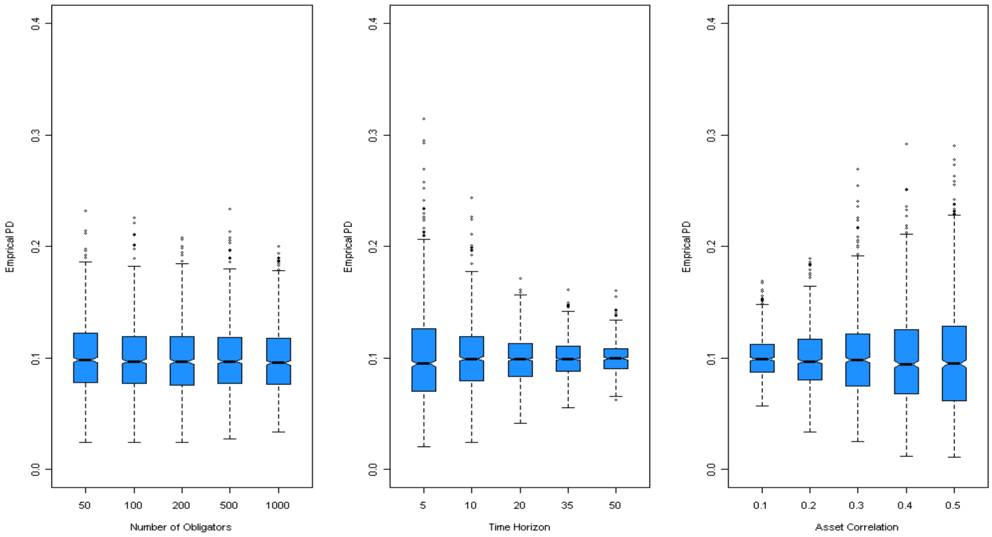


Figure 4. The Empirical PD vs. Portfolio Size (the left, holding other setting as the benchmark setting), Tracked time horizon (the middle, holding other setting as the benchmark setting), the Asset correlation (the right, holding other setting as the benchmark setting). The true value of PD = 0.10, Empirical PD are calculated based on 1000 repeated simulated data sets,

Table 4. Empirical PD (True Value of PD = 0.10)

		Mean	**Median**	**SD**	**Skewness**	**Kurtosis**	**[P5 , P95]**
	50	0.102	0.098	0.033	0.528	3.200	[0.054 , 0.162]
	100	0.100	0.097	0.033	0.661	3.489	[0.054 , 0.159]
N	*200*	0.099	0.096	0.032	0.484	3.025	[0.053 , 0.157]
	500	0.100	0.097	0.031	0.661	3.592	[0.056 , 0.159]
	1000	0.099	0.096	0.030	0.512	3.026	[0.055 , 0.153]
	5	0.101	0.095	0.043	0.980	4.484	[0.045 , 0.186]
	10	0.101	0.099	0.031	0.645	3.917	[0.054 , 0.156]
T	*20*	0.100	0.099	0.021	0.272	2.771	[0.066 , 0.137]
	35	0.100	0.099	0.016	0.271	2.857	[0.076 , 0.127]
	50	0.100	0.099	0.014	0.321	3.325	[0.078 , 0.123]
	0.1	0.101	0.099	0.018	0.448	3.300	[0.072 , 0.132]
	0.2	0.099	0.097	0.026	0.355	3.031	[0.059 , 0.143]
ρ	*0.3*	0.101	0.098	0.035	0.791	4.306	[0.052 , 0.159]
	0.4	0.100	0.094	0.041	0.732	3.621	[0.044 , 0.173]
	0.5	0.100	0.095	0.049	0.765	3.495	[0.035 , 0.192]

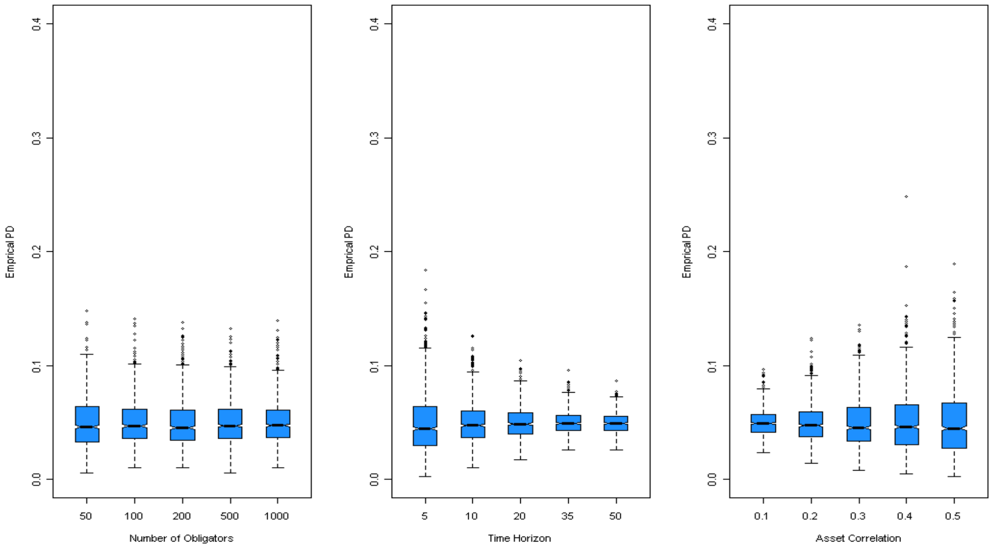


Figure 5. The Empirical PD vs. Portfolio Size (the left, holding other setting as the benchmark setting), Tracked time horizon (the middle, holding other setting as the benchmark setting), the Asset correlation (the right, holding other setting as the benchmark setting). The true value of PD = 0.05, Empirical PD are calculated based on 1000 repeated simulated data sets,

Table 5. Empirical PD (True Value of PD = 0.05)

		Mean	Median	SD	Skewness	Kurtosis	[P95 , P5]
	50	0.050	0.046	0.022	0.673	3.522	[0.020 , 0.088]
	100	0.050	0.047	0.020	0.785	3.938	[0.023 , 0.086]
N	*200*	0.050	0.046	0.021	0.937	4.030	[0.023 , 0.088]
	500	0.050	0.047	0.020	0.788	3.626	[0.024 , 0.088]
	1000	0.050	0.048	0.019	0.914	4.451	[0.024 , 0.085]
	5	0.050	0.044	0.027	1.101	4.464	[0.016 , 0.105]
	10	0.050	0.047	0.019	0.844	3.834	[0.024 , 0.087]
T	*20*	0.050	0.049	0.014	0.561	3.324	[0.030 , 0.074]
	35	0.050	0.049	0.010	0.562	3.549	[0.035 , 0.069]
	50	0.050	0.049	0.009	0.379	3.308	[0.036 , 0.064]
	0.1	0.050	0.049	0.011	0.488	3.411	[0.033 , 0.070]
	0.2	0.050	0.048	0.017	0.791	3.866	[0.027 , 0.082]
ρ	*0.3*	0.050	0.046	0.022	0.777	3.464	[0.020 , 0.091]
	0.4	0.050	0.046	0.027	1.321	7.083	[0.016 , 0.097]
	0.5	0.050	0.045	0.029	0.937	4.082	[0.012 , 0.104]

IV-3. Estimated Results of the Asset Correlation

In this section, we present the estimated results of the asset correlation following the procedure described in ***III-1***. Firstly, we present the result based on the benchmark exercise data, in order to show in details the working procedure. Secondly, we present the estimated results based on 50 simulated data sets, so that we can get more insights in the general performance of the different solutions.

IV-3-1. Result Based on the Benchmark Exercise Data

There are two ways to model the asset correlation: Tarashev solution and LMM solution. Both solutions employ a Bayesian framework for inference. Before presenting the result of Bayesian estimation for both solutions, we can first take a look at the estimates calculated by the non-Bayesian method.

More specifically, the non-Bayesian estimates presented here are: (1) the point estimate of the asset correlation calculated as (F3-1), and (2) the estimate calculated by the LMM model (F3-3) with ML estimation. The uncertainty around such estimates is derived by a bootstrap scheme. The bootstrap scheme is designed as the following:

> *Given the asset return data $\boldsymbol{X}_{100\times 120}$, we randomly simulate 120 columns out of the original 120 columns of $\boldsymbol{X}$ with replacement, so that the bootstrap data $\boldsymbol{X}_{boot}$ is of the same dimension as the original data. Repeat such procedure 1000 times, and for each $\boldsymbol{X}_{boot}^{[i]}$ $(i \in \{1,2,\dots,1000\})$, an estimate of the asset correlation can be calculated. The uncertainty of the estimate can be portrayed as the distribution of these 1000 values.*

The result is presented graphically in Figure 6 and numerically in Table 6. One can see that these two non-Bayesian estimates are both around 0.22 while the true value of the asset correlation equals to 0.25, and the standard deviations of estimates based on 1000 bootstrap data sets are both around 0.02. This is close to what we find under the LMM solutions (see the next paragraph).

Table 6. Non-Bayesian Estimation of the Asset Correlation

	Estimates	**Bootstrap**			
		Mean	*S.D.*	*Skewness*	*Kurtosis*
Point Estimation	0.2216	0.222	0.020	0.079	2.935
LMM Estimation	0.2206	0.218	0.021	0.019	2.877

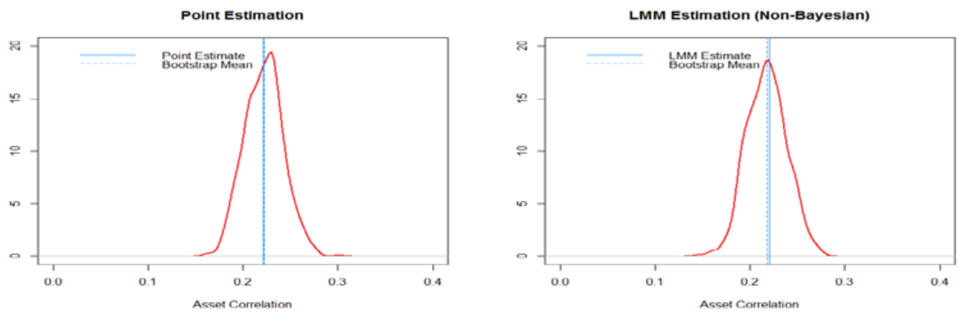


Figure 6. The distribution of 1000 Bootstrap estimates by the point estimation method (left plot) and the LMM mode l(right plot). The solid blue line stands for the estimates based on original data, and the dash blue line stands for the mean of bootstrap estimates.

Table 7. Bayesian Estimation of the Asset Correlation

Solution	Mean	Median	S.D.	Skewness	Kurtosis	MC-Error
LMM	0.2243	0.2230	0.0240	0.2909	3.7609	0.0003
Tarashev	0.2566	0.2454	0.0837	0.6626	3.3210	0.0009

** The descriptive statistics is calculated based on 10000 iterations of MCMC sample with 5000 burn-in..*

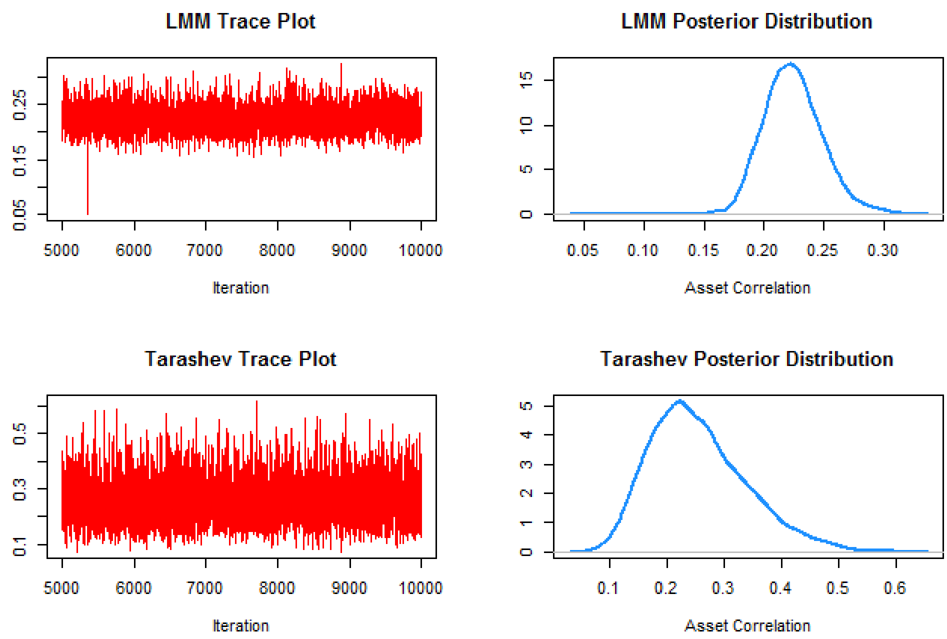


Figure 7. Posterior Distribution derived by a LMM model (the top row) and the Tarashev solution (the bottom row). Left column of the plots is about MCMC samples vs. iteration (excluding the first 5000 burn-in), and the right column are the density plots of the posteriori.

Now, we consider the Bayesian estimation of the asset correlation. The Bayesian estimation in this paper is all run by WinBUGS 1.4 with the Gibb sampler and the estimated results of two solutions are presented in Table 7 and Figure 7.

From the trace plots in Figure 7, both MCMC chains have adequately good mixing rate. We take the posterior mean as the estimate of the asset correlation. The convergence of such mean estimate can be assured by the Geweke diagnostic with the first 10% fraction of the chain vs. the last 50% fraction, which returns Z-scores with small absolute value: 0.4821 (for the LMM estimate) and -0.3095 (for the Tarashev estimate). Based on Table 7, the Tarashev solution gives an over-estimated asset correlation with a departure of 0.0066 to the true value (i.e. 0.25), while the LMM model gives an under-estimated value with a departure of 0.0257 to the true value. However, the posterior variance of the LMM estimate is smaller than the Tarashev estimate. Besides, both posteriori, given the benchmark exercise data, is approximate to a normal distribution, but the Tarashev posterior is more skewed.

IV-3-2. Result Based on Repeatedly Simulated Data

Now, we run both solutions of the asset correlation with 50 repeatedly simulated asset return data, and these 50 data sets are simulated with the benchmark setting (i.e. *N=100, T=10, ρ=0.25, PD=0.1, μ=6, σ=15.*). For each simulated data set ($\mathbf{X}_{100\times 120}^{[\mathbf{b}]}, b \in \{1,2,\dots,50\}$), a posteriori of the asset correlation can be derived, and Table 8 shows the descriptive statistics of some featured posterior characteristics, under the LMM or Tarashev solution. To quantify the departure of the posterior mean to the true value of ρ, we use the sum of square departure (SSD) calculated as $SSD = \sum_b (\hat{\rho}_b - \rho)^2$.

Table 8. Descriptive Statistics of Posterior Distributions of the Asset Correlation based on 50 simulated data sets

	Mean	**P5**	**P95**	**SSD**
LMM				
Posterior Mean	0.2538	0.2102	0.2972	0.0394
Kurtosis of Posteriori	3.5949	3.1297	4.0710	
Skewness of Posteriori	0.2686	0.1770	0.3689	
S.D. of Posteriori	0.0255	0.0226	0.0281	
Tarashev				
Posterior Mean	0.2833	0.2424	0.3239	0.0865
Kurtosis of Posteriori	3.0834	2.8420	3.3637	
Skewness of Posteriori	0.5292	0.4020	0.6686	
S.D. of Posteriori	0.0872	0.0818	0.0919	

* The posterior estimate is derived with Gibbs sampling with 10000 iterations and 5000 burn-in.

One can see that, both estimates tend to overestimate the asset correlation. The posterior distributions derived under both solutions all show slight skewness, and in the case of LMM solution, the posteriori has larger kurtosis. Tarashev estimate of the asset correlation has a bigger SSD and the standard deviation of the posteriori is averagely around 0.09. In other words, LMM estimate of the asset correlation is more accurate, and the uncertainty around the posterior estimate is also smaller. Besides, Figure 8 shows that most of the posterior estimates have small absolute Geweke Z-score (rule of thumb: < 2), which indicates that the posterior mean is converged. However, it is important to note that the LMM solution carries more computational burden. For a single run of the estimation with $\mathbf{X}^{[\mathbf{b}]}_{100\times120}$, the LMM solution take averagely 10 times longer than the Tarashev solution in a computer with Intel-i5 CPU 2.27 GHz and 4 GB RAM. In practice, one may choose to tolerate the estimation bias for the fast calculation.

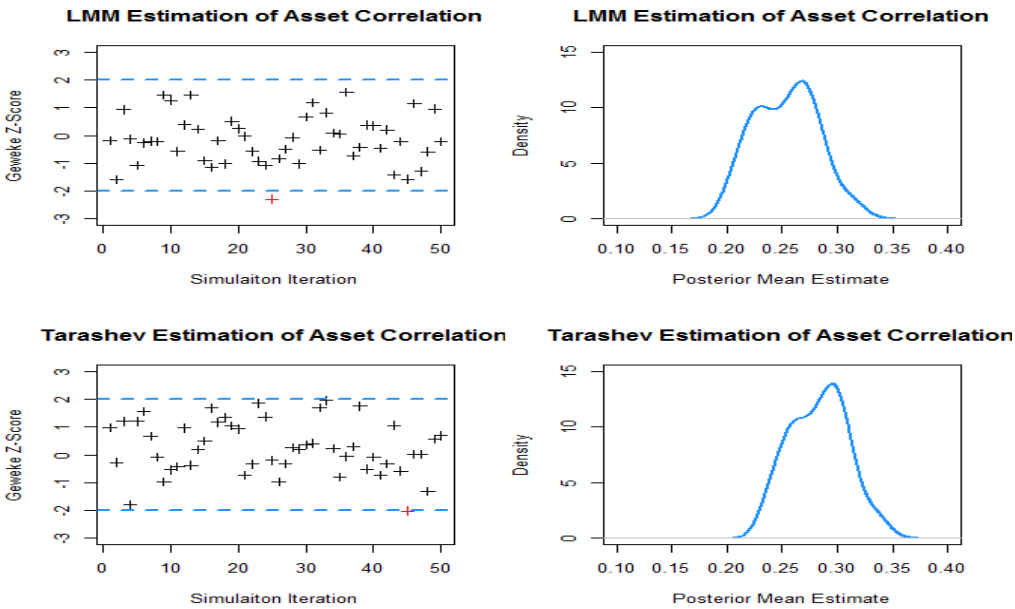


Figure 8. Left Row shows the Geweke Z-scores of MCMC estimate based on 50 simulated data sets under the LMM and Tarashev solutions respectively; Right Row shows the density plot of the 50 posterior means of the asset correlation under the LMM and Tarashev solutions respectively, and the estimates with $|Z\ Score| > 2$ *are excluded.*

IV-4. Estimated Results of PD

In this section, we present the estimated results of PD in the same manner as the previous section.

IV-4-1. Result Based on the Benchmark Exercise Data

As discussed in ***III-2*** (last two paragraphs), the estimate of the unconditional PD can be derived in two ways in the GLMM framework, dubbed as the Formal and Alternative estimate of PD respectively. The inference result of these two estimates of PD, given different priori for the asset correlation, is presented in Table 9~10 and Figure 9~10.

Table 9. Bayesian Inference for the Formal Estimate of PD

Prior of ρ	**Mean**	**Median**	**S.D.**	**Skewness**	**Kurtosis**	**MC-Error**	**Z-Score**
LMM	0.1147	0.1153	0.0220	0.0076	4.7190	0.0011	-0.2759
Tarashev	0.1182	0.1166	0.0253	0.4620	3.5793	0.0034	1.0340

** The descriptive statistics is calculated based on 10000 iterations of MCMC sample and burned the first 5000; Z-score is the result of the Geweke diagnostic with the first 10% fraction of the chain vs. the last 50% fraction.*

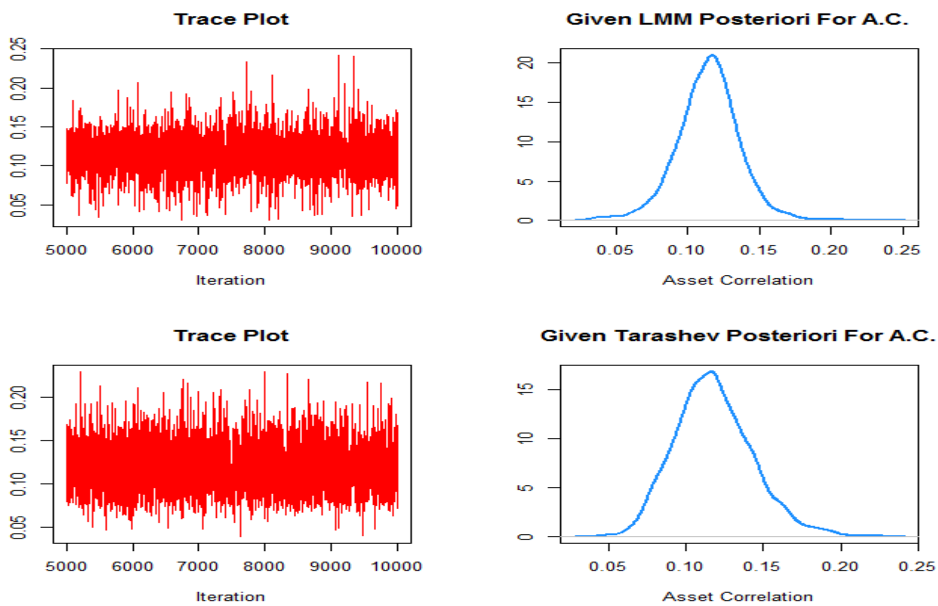


Figure 9. Posterior Distribution of Formal Estimate of PD given different priori for the asset correlation, derived by a LMM model (the top row) and the Tarashev solution (the bottom row) respectively. Left column of the plots is about MCMC samples vs. iteration (excluding the first 5000 burn-in), and the right column are density plots of the posteriori.

Table 10. Bayesian Inference for the Alternative Estimate of PD

Prior of ρ	Mean	Median	S.D.	Skewness	Kurtosis	MC-Error	Z-Score
LMM	0.0942	0.0940	0.0090	0.1185	3.0912	0.0001	-0.7529
Tarashev	0.0941	0.0939	0.0092	0.1557	3.0053	0.0001	0.3542

* *The descriptive statistics is calculated based on 10000 iterations of MCMC sample and burned the first 5000; Z-score is the result of the Geweke diagnostic with the first 10% fraction of the chain vs. the last 50% fraction,*

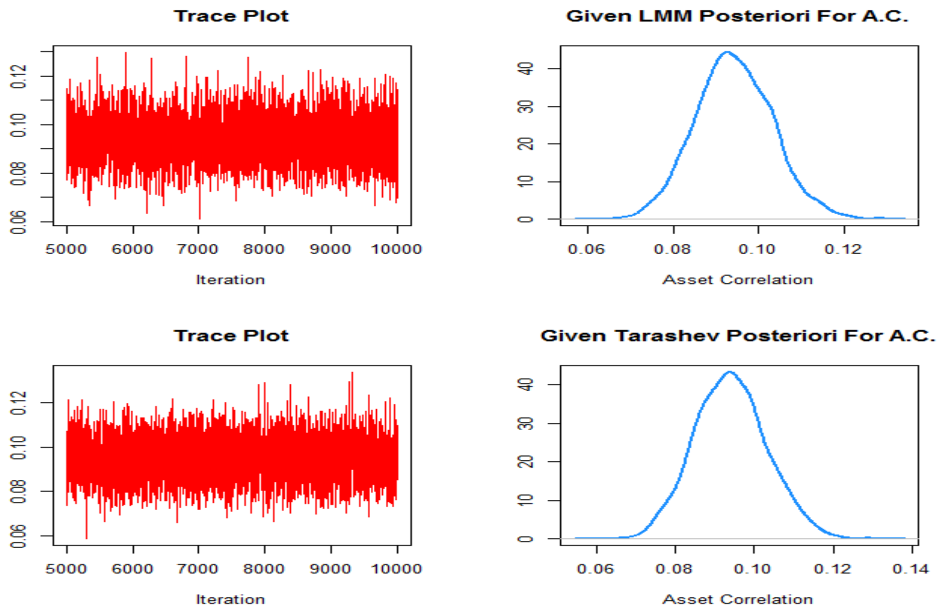


Figure 10. Posterior Distribution of Alternative Estimate of PD given different priori for the asset correlation, derived by a LMM model (the top row) and the Tarashev solution (the bottom row) respectively. Left column of the plots is about MCMC samples vs. iteration (excluding the first 5000 burn-in), and the right column are density plots of the posteriori.

From the trace plots in Figure 9~10, all MCMC chains have adequately good mixing rate and small absolute Z-scores indicate that the posterior means are converged. We take the posterior means as the estimates of the asset correlation. They have very small difference between the occasions given LMM or Tarashev posteriori for the asset correlation. This holds for both formulations of PD estimate (i.e. formal and alternative). Therefore, a more interesting aspect of the above should be the comparison between the Formal and Alternative estimate of PD.

Regardless the given prior for the asset correlation, the Formal estimates of PD have an average 0.0165 departure from the true value (i.e. 0.10), while the Alternative estimates have an average 0.0058 departure. The MC-errors in the Alternative formulation is over 10 times smaller than the MC-errors in the Formal formulation, In other words, the accuracy of the posterior means in the case of the Alternative formulation is higher. Besides, the standard deviation of the Alternative posteriori is less than half of the standard deviation of the Formal posteriori. To sum up, based on the benchmark exercise data, the Alternative estimate of PD performs better.

Now we take a look at the relationship between the posterior distributions of PD and the asset correlation (2D plots with contour lines in Figure 11 and 3D plots in Appendix C). Based on the bottom row of Figure 11, the posteriori of PD and ρ have very slight correlation when PD is estimated in the Alternative formulation, approximately 0.02 given either LMM posteriori or Tarashev posteriori as the priori for the asset correlation. On the other hand, when PD estimated is derived in the Formal formulation, the posteriori of PD and ρ has 0.137 correlation given LMM posteriori for the asset correlation. This correction increases to 0.532, given Tarashev posteriori for the asset correlation.

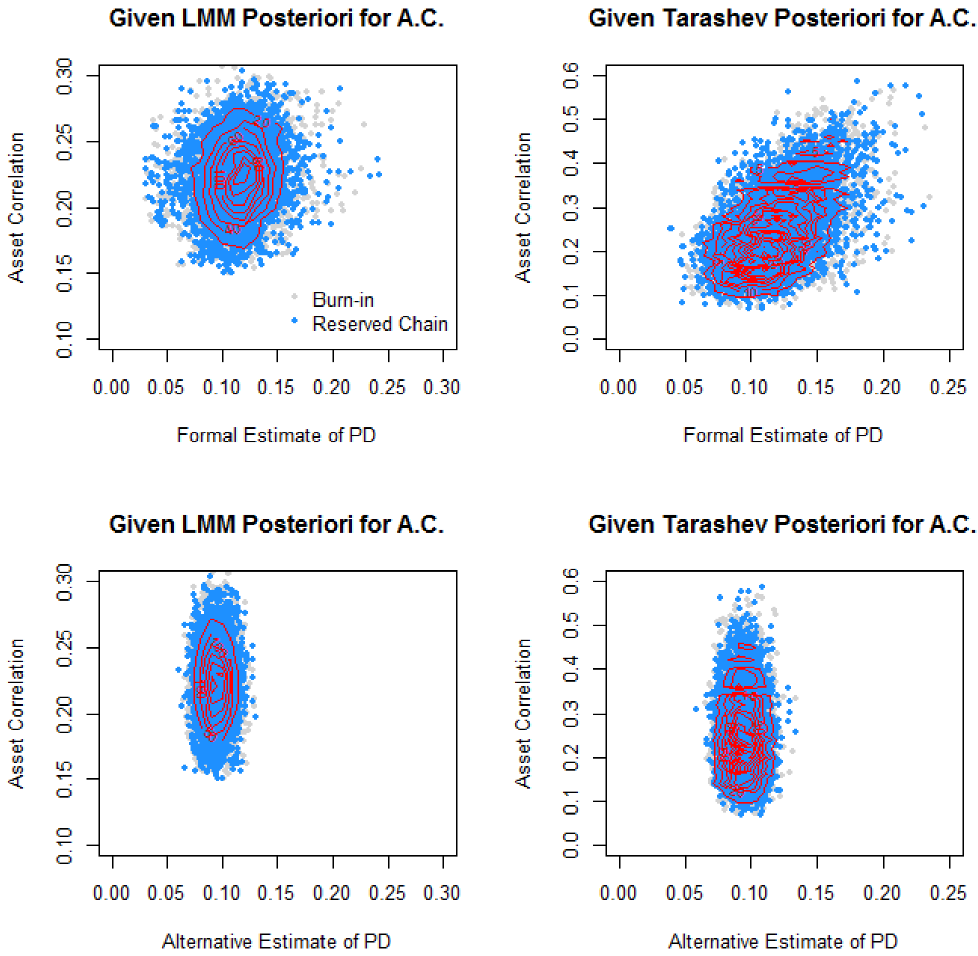


Figure 11. Scatterplot with contour lines of posterior PD estimate (Formal estimate in the top row and Alternative estimate in the bottom row) and posterior asset correlation, given different priori for the asset correlation derived by a LMM model (left column) and the Tarashev solution (right column); The grey points stands for the first 5000 burned MC chain, while the blue points stands for the reserved chain.

IV-3-2. Result Based on Repeatedly Simulated Data

In this subsection, we run the estimation procedure of PD (i.e. Formal estimate, Alternative estimate given either LMM posteriori or Tarashev posteriori for the prior of the asset correlation) over 50 simulated default data sets ($\mathbf{I}_{1\times 10}^{[b]}, b \in [1,2,\ldots,50]$). These data sets are consistent with the simulated asset return matrixes ($\mathbf{X}_{100\times 120}^{[b]}$) in ***IV-3-2***. The results are also presented in a similar manner in Table 11~12 and Figure 12~13.

Based on Table 11~12, no matter deriving the PD estimation in the Formal or Alternative way, the choice between LMM or Tarashev posteriori for the asset correlation shows only small difference across all the posteriori characteristics. The interests will then be put on the comparison between the performance of Formal and Alternative estimate of PD.

Starting with comparison across some featured posteriori characteristics, the Formal estimate of PD produce in general more peaked and skewed posteriori, and the standard deviation of the posterior distribution is averagely around 4 times larger than the cases of the Alternative estimate. One more highlight, the posterior Alternative PD estimate has trivial correlation with posterior ρ correlation across the 50 simulated data sets. Actually, when a non-informative prior is assigned to ρ, the posterior distribution of the Alternative PD estimate barely changes. This formulation is only related to the latent common factor. On the other hand, the posterior Formal PD estimate has significant non-zero correlation with posterior ρ. Especially when Tarashev posteriori is used as the prior of ρ, the correlation is averagely around 0.4 and has a 95% confidence interval of [0.22, 0.58].

Now, we focus on the comparison between the posterior means in all occasion, since it is usually taken as the representative posterior estimate. First, based on the left rows of Figure 12 ~13, most of the posterior means are assured to be converged by the Geweke diagnostics with the first 10% fraction of the chain vs. the last 50% fraction. Besides, the Z-scores of the Formal PD estimates are distributed closer around zero, which indicates relative higher accuracy of the posterior mean. This indication is confirmed by the variance of the 50 estimates of posterior mean across simulated data sets, based on the right rows of Figure 12 ~13. However, the SSD shows that the both estimate of PD has a similar scale of bias in general. But in the case of the Alternative estimate of PD, the mean of the 50 posterior means are closer to the true value of PD (i.e. 0.1). Last thing to notice, deriving PD estimate in the Alternative way costs only half of the time needed in the Formal way, and is free of some computational dilemma in WinBUGS.

Implementation Note:

As discussed in ***III-2***, both PD estimates are derived with a Probit GLMM. Due to PD is usually a very small value, this raises some computational issues, which are well recognized when the estimation procedure is implemented by WinBUGS (see WinBUGS User Manual). The MCMC algorithm tends to break down under the Formal formulation of PD estimate, even with less diffused prior of u and other compromised parameterization, while the Alternative formulation is free of this dilemma.

Table 11. Descriptive Statistics of Posterior Distributions of PD
based on 50 simulated default data sets, Given LMM posteriori for A.C.

	Mean	**P5**	**P95**	**SSD**
Formal Estimate of PD				
Posterior Mean	0.1121	0.0697	0.1636	0.0466
Kurtosis of Posteriori	5.1696	3.5371	10.1843	
Skewness of Posteriori	0.7198	0.1785	1.5305	
S.D. of Posteriori	0.0345	0.0195	0.0535	
Correlation with ρ	0.1201	0.0181	0.2004	
Alternative Estimate of PD				
Posterior Mean	0.1031	0.0601	0.1632	0.0505
Kurtosis of Posteriori	3.0421	2.9178	3.1580	
Skewness of Posteriori	0.1682	0.0785	0.2444	
S.D. of Posteriori	0.0090	0.0074	0.0106	
Correlation with ρ	0.0030	-0.0171	0.0210	

* The posterior estimate is derived with Gibbs sampling with 10000 iterations and 5000 burn-in.

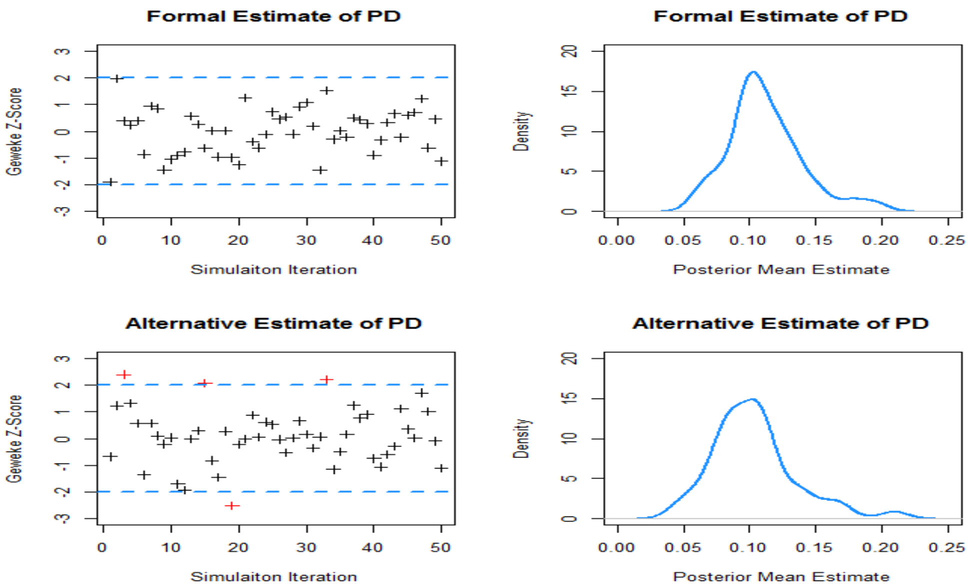


Figure 12. Plots about the Posterior Mean given LMM posteriori for the asset correlation. Left Row shows the Geweke Z-scores of MCMC estimate based on 50 simulated data sets under the Formal and Alternative formulation of PD estimation respectively; Right Row shows the density plot of the 50 posterior means of the Formal and Alternative estimate of PD respectively, and the estimates with $|Z\ Score| > 2$ *are excluded.*

Table 12. Descriptive Statistics of Posterior Distributions of PD based on 50 simulated default data sets, Given Tarashev posteriori for A.C.

	Mean	P5	P95	SSD
Formal Estimate of PD				
Posterior Mean	0.1184	0.0790	0.1707	0.0519
Kurtosis of Posteriori	4.7591	3.2957	7.5151	
Skewness of Posteriori	0.7078	0.2511	1.4234	
S.D. of Posteriori	0.0383	0.0231	0.0567	
Correlation with ρ	0.3983	0.2240	0.5811	
Alternative Estimate of PD				
Posterior Mean	0.1074	0.0599	0.1639	0.0505
Kurtosis of Posteriori	3.0186	2.8860	3.1339	
Skewness of Posteriori	0.1611	0.1073	0.2529	
S.D. of Posteriori	0.0092	0.0074	0.0108	
Correlation with ρ	-0.0030	-0.0181	0.0174	

* The posterior estimate is derived with Gibbs sampling with 10000 iterations and 5000 burn-in.

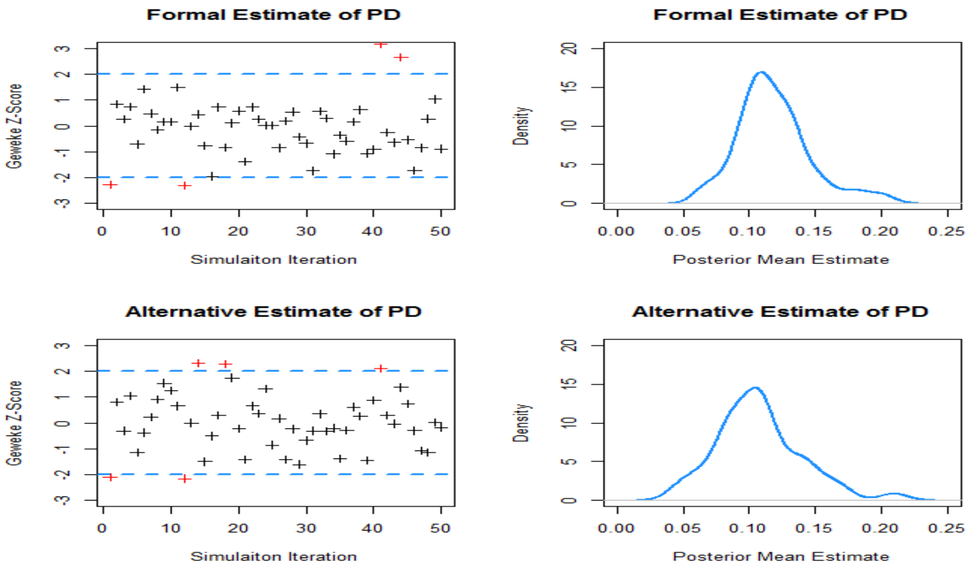


Figure 13. Plots about the Posterior Mean given Tarashev posteriori for the asset correlation. Left Row shows the Geweke Z-scores of MCMC estimate based on 50 simulated data sets under the Formal and Alternative formulation of PD estimation respectively; Right Row shows the density plot of the 50 posterior means of the Formal and Alternative estimate of PD respectively, and the estimates with $|Z\ Score| > 2$ *are excluded.*

IV-5. Estimated Results of VaR

Quantifying the influence of the estimation uncertainty on the VaR is the most important task in this paper. In this section, we first show the add-on of VaR based on the benchmark exercise data to emulate a specific practical exercise. Then, we run a simulation-based exercise to demonstrate the influence of estimation uncertainty to the VaR add-on from several aspects.

IV-5-1. Result of VaR Add–on Based on the Benchmark Exercise Data

First, we summarize the input of VaR calculation up to now. Based on the benchmark exercise data, we have known the true value of PD and ρ is 0.1 and 0.25 respectively. There are diverse estimate of PD and ρ, which can be listed as the following (Table 13)

Table 13. Summary of the PD and ρ Estimation

Case		Estimation of ρ			Estimation of PD	
		Posterior Mean	S.D. of Posteriori		Posterior Mean	S.D. of Posteriori
1	*LMM*	0.2243	0.0240	*Formal*	0.1147	0.0220
2	*LMM*	0.2243	0.0240	*Alternative*	0.0942	0.0090
3	*Tarashev*	0.2566	0.0837	*Formal*	0.1182	0.0253
4	*Tarashev*	0.2566	0.0837	*Alternative*	0.0941	0.0092

Taking the posterior mean as the point estimate, the Naive VaR can be calculated based on (F3-8) and correct VaR can be derived following the procedure in ***III-3***. The result is presented in Table 14.

Table 14. VaR Add-on based on the Benchmark Exercise Data

Case	Naive VaR	Add-on		
		Noise in $\hat{\rho}$ *only*	*Noise in* $\widehat{PD}$ *only*	*Noise in both*
1	0.4576	-0.0027	0.0099	0.0104
2	0.4042	0.0515	0.0029	0.0029
3	0.4974	-0.0005	-0.0209	0.0455
4	0.4366	-0.0583	-0.0310	0.0037

Based on the result of the Naïve VaR, using different estimate of ρ and PD can cause quite significant difference. When the estimation uncertainty of both parameters are taken into account, we will have a positive VaR add-on, and these add-on can take up approximately 0.75 ~ 9% of the Naïve VaR.

IV-5-2. Result of VaR Add–on Based on a Simulation Exercise

The simulation exercise is designed as the following:

1. Assume that the estimate of the asset correlation and PD are $\hat{\rho}$ and $\widehat{PD}$ respectively, and they are both given;
2. The uncertainty around the estimate $\hat{\rho}$ is assumed to be well portrayed by $Beta\,(\alpha, \beta)$, and the parameters (α, β) are determined by two equations: (1) $\alpha/(\alpha+\beta) = \hat{\rho}$ and (2) $\sqrt{\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)} = sd(\hat{\rho})$, where $sd(\hat{\rho})$ is also known;
3. The uncertainty around the estimate $\widehat{PD}$ is assumed to be well portrayed by $N\big(\,\widehat{PD}\,,\ (0.01+0.07\hat{\rho})^2\,\big)$

Following such design, we have determined the estimate $(\widehat{PD}, \hat{\rho}\,)$ and their marginal and joint distribtuions. Then, the naive VaR can be calculated based on (F3-8), and we can follow the procedure in ***III-3*** to derive the correct VaR. There is one remark warranted for the above simulation. Using a normal distribution for the estimate $\widehat{PD}$ is inspired by the Empirical PD exercise in ***IV-2***. The formulation of the s.d. of this normal distribution is based on a regression model. Taking the S.D. data given diverse ρ setting in Table 4 (i.e. simulation given true value of PD equals to 0.1) as the input for the response variable, a simply linear regression returns the $S.D. = 0.01 + 0.07\hat{\rho}$ with $R^2 = 0.99$. Although the coefficient of $\hat{\rho}$ will change given input of S.D (say Table 5, given true value of PD equals to 0.05), it is sufficient for illustrative purpose. Based on the Figure 14, the hypothetical normal distribution $N\big(\,\widehat{PD}\,,\ (0.01+0.07\hat{\rho})^2\,\big)$ can well mimic the distribution of Empirical PD, under the benchmark setting while given free setting for ρ.

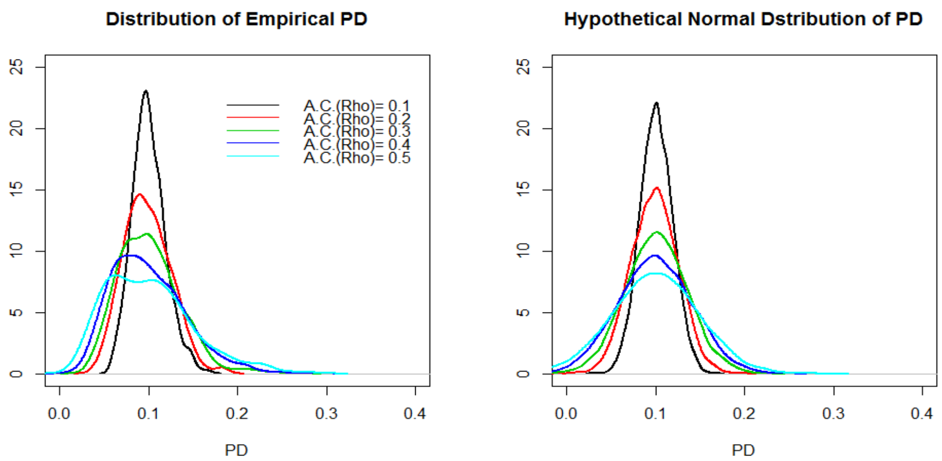


Figure 14. Left plot shows the density plots of the Empirical PD given the benchmark exercise setting while given diverse values for ρ*; Right plot shows the densidty plots of* $N(\,0.1\,, (0.01+0.07\hat{\rho})^2\,)$ *given diverse values for* ρ.

For the simulation, we use the candidate setting as the following.

$$\hat{\rho} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$$
$$\widehat{PD} \in \{0.05, 0.1, 0.15, 0.2, 0.25\}$$
$$sd(\hat{\rho}) \in \{0.1, 0.15, 0.2, 0.25, 0.3\}$$

The detailed simulation scheme for deriving VaR add-on can be expressed as in Table 15. To note, this simulation scheme taking into account the uncertainty of both $\hat{\rho}$ and $\widehat{PD}$. In order to have more comprehensive look, we consider two extra scenarios: noise in $\hat{\rho}$ only and noise in $\widehat{PD}$ only. The simulation schemes of these two scenarios are the simplified versions of scheme for the noise in both, and therefore will not be specified here.

Table 15. Simulation Scheme (II)

Step	Instruction
1	Setting $\hat{\rho} = 0.1$ and $sd(\hat{\rho}) = 0.1$, and simulate *B* values of $\hat{\rho}$ from a Beta distribution (denote as $\hat{\rho}_{[b]}, b \in \{1,2,\dots,B\}$), where the parameters of the Beta distribution is calculated in the way presented in the previous context of IV-5-2;
2	Setting $\widehat{PD} = 0.05$, and calculate the Naïve VaR. For each $\hat{\rho}_{[b]}$, simulate *K* values $\widehat{PD}_{[b,k]}$ from a normal distribution with mean 0.05 and s.d.=$0.01 + 0.07\hat{\rho}$. If the simulated $\widehat{PD}_{[b,k]}$is negative, it is abandoned in later calculation;
3	Simulate *K* values for common factor M from $N(0,1)$. Given also $\hat{\rho}_{[b]}$ and $\widehat{PD}_{[b,k]}$, calculate the *K values of* $E(L_N \| M_t, PD, \rho)$ as in (F3-9);
4	Using simulated values of $E(L_N \| M_t, PD, \rho)$ to form the loss distribution, and then calculate the 99% quantile, which is the correct VaR. Calculate add-on;
5	Repeat Step-2 ~3 by replace $\widehat{PD}$ with all candidate values from $\{0.05, 0.1, 0.15, 0.2, 0.25\}$;
6	Repeat Step-1~4 by replace ($\hat{\rho}$, $sd(\hat{\rho})$) with all possible combination based on their candidate values;
7	Repeat Step 1~6 up to 50 times, and in turn derive the mean and 95% CI for each estimate of VaR add-on.

** B and K are set large enough to guarantee that the number of total simulated loss values is larger than 100000.*

(1) Considering Only the Uncertainty of ρ Estimate

First of all, we consider the scenario that noise in $\hat{\rho}$ only. The result can be presented in Figure 15, and more detailed numerical result is presented in Table 16 in Appendix D. One can see that, the uncertainty of $\hat{\rho}$ causes a positive add-on, and averagely the add-on tends to increase when the uncertainty grows (i.e. $\mathrm{sd}(\hat{\rho})$ increases).

Moreover, when the PD estimate is larger than 0.2, the add-on tends to decrease as $\hat{\rho}$ increases. In plain terms, when the investors expect the general default possibility is higher than 0.2 and also are aware that the obligators in the portfolio share weak financial connection, the correct VaR will suggest to reserve more economic capital than the naïve VaR for hedging default risk. If the obligators have high asset dependency, the investor may reserve less, since companies have higher possibility to survival as a union than individuals through tough market.

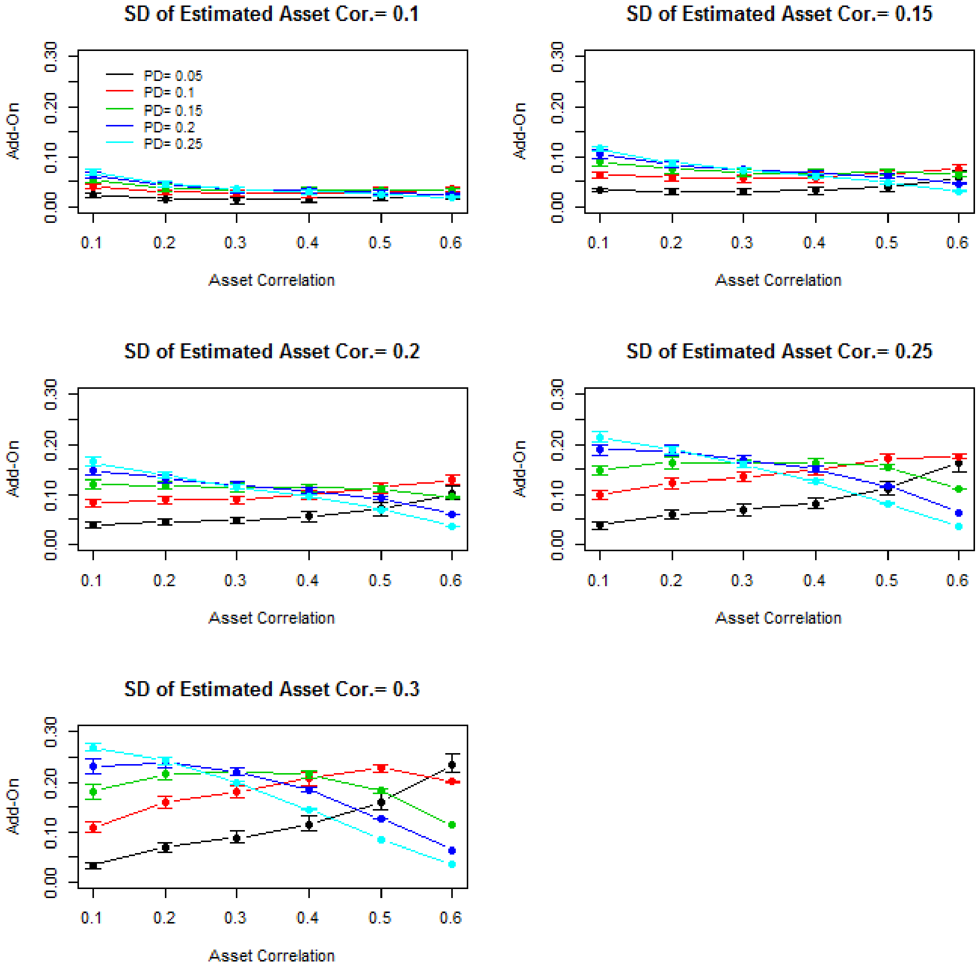


Figure 15. Plot of VaR Add-on in the scenario of Noise in $\hat{\rho}$ only, given different setting of $\widehat{PD}, \hat{\rho}$ and $sd(\hat{\rho})$; the points represents the mean of corresponding Add-on estimate over 50 repeated simulation, and with 95% CI around.

On the other hand, when PD is believed to be smaller than 0.1, the add-on tends to be positively related to $\hat{\rho}$. This indicates that, when the general PD is expected to be low, the correct VaR will call to reserve more capital if the obligators in the portfolio share high financial connection. This can be interpreted as the preparation for a hinged hit. To elaborate this opinion, when the obligators are expected to be performing well individually and also have tight business connection, the risk manager may want to prepare for the worst scenario: for instance, one of the obligator has bankrupt and this shock will significantly injure all the other obligators due to their commercial collaboration.

These trends get more and more obvious while the uncertainty of $\hat{\rho}$ grows. One may find such operation paradoxical. To note, when the uncertainty of $\hat{\rho}$ is small (for instance,

$sd(\hat{\rho}) = 0.1$), the correct VaR suggest similar add-on when the asset correlation is higher than 0.3, no matter what the estimated value of PD is, and suggest higher add-on when asset correlation is low and PD is expected to be high. The trends discussed in the previous two paragraphs mainly appear when $\hat{\rho}$ is highly uncertain (i.e $sd(\hat{\rho}) > 0.2$).

(2) Considering Only the Uncertainty of PD Estimate

Now, we consider the scenario that noise in $\widehat{PD}$ only. The result can be presented in Figure 16, and more detailed numerical result is presented in Table 17 in Appendix D. One can see that, the uncertainty of $\widehat{PD}$ has generally small impact on the VaR add on. However, due to the simulation design, the uncertainty of $\widehat{PD}$ is positively related to $\hat{\rho}$. This can explain why the 95% CI round the add-on estimates grows wider as $\hat{\rho}$ increasing.

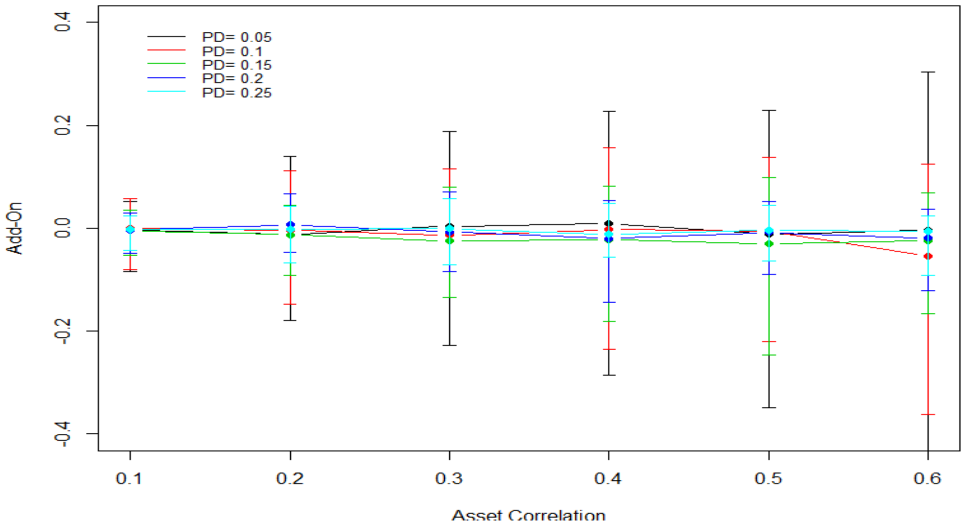


Figure 16. Plot of VaR Add-on in the scenario of Noise in $\widehat{PD}$ only, given different setting of $\widehat{PD}$and $\hat{\rho}$; the points represents the mean of corresponding Add-on estimate over 50 repeated simulation, and with 95% CI around.

(2) Considering the Uncertainty of both ρ and PD Estimate

Finally, we consider the scenario that noise in both estimates. The result can be presented in Figure 17, and more detailed numerical result is presented in Table 18 in Appendix D.

The add-ons are positive in general. The features in Figure 17 are similar to Figure 15. However, one can see that the correct VaR suggests to reserve more economic capital than in Figure 15, when PD is expected to be small (i.e. 0.05) and the asset correlation is high. This is due to the uncertain of $\widehat{PD}$ increases along with $\hat{\rho}$. Therefore, it is rational to reserve more capital when the investors are not certain about the PD estimate but aware of the asset correlation is high.

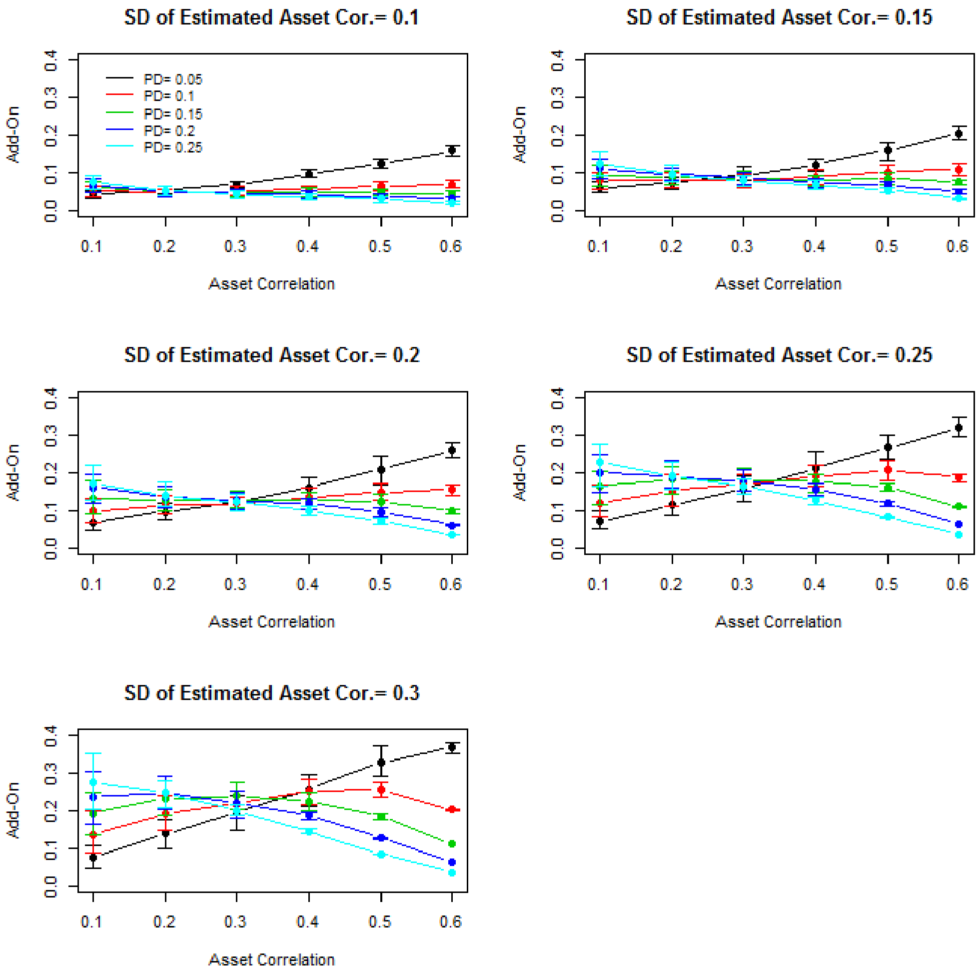


Figure 17. Plot of VaR Add-on in the scenario of Noise in both $\hat{\rho}$ and $\widehat{PD}$, given different setting of $\widehat{PD}, \hat{\rho}$ and $sd(\hat{\rho})$; the points represents the mean of corresponding Add-on estimate over 50 repeated simulation, and with 95% CI around.

Chapter V

Conclusion

In the previous chapter, we confirm that the estimation uncertainty of the asset correlation and PD and its impact on measuring VaR.

For the estimation of asset correlation, with 50 simulated data sets, we witness that the uncertainty around the point estimate is ineligible (around 0.026 in LMM solution and 0.087 in Tarashev solution). The LMM solution tends to be more accurate based on the SSD, while it has a heavier load in calculation than the Tarashev solution. In practice, the balance between the accuracy and computational efficiency is left for individual appreciation.

For the estimation of PD, we can formulate the inference procedure in two way based on (F3-7) and (F3-5) respectively, dubbed as the Formal estimate and the Alternative estimate. With 50 simulated data sets, we witness only small changes in the posteriori of PD given LMM or Tarashev posteriori as the prior for the asset correlation. However, in the former case, the performance is slightly better based on SSD. The more interests are then invested on the comparison between the performance of the Formal and Alternative estimates. Given the LMM prior for ρ, the Alternative estimate is less accurate than the Formal estimate, while the conclusion turns over when given the Tarashev prior for ρ. However, comparing to the Formal estimate of PD, the posterior Alternative estimate has generally (a) around 4 times smaller variance, (b) better approximation to a normal distribution, and (c) almost no correlation with the posterior ρ. The last finding inspires us to try a diffuse prior for ρ, and the Alternative estimate of PD stays quite similar. This confirms that it is robust to the prior of ρ. One can find in Figure 12 and 13 that the distribution of the Formal estimate (posterior mean) is more concentrated, based on the 50 simulation runs, which might due to the fact that it takes advantages of extra information about the asset correlation.

For the VaR add-on, we first show the result based on the benchmark exercise data. Most of the add-ons are positive, and can take up around 0.75~9% of the naïve VaR. However, such result is based on one single data sets with diverse way of specifying the joint distribution of (PD, ρ). Later, we further investigate the influence of the estimation uncertainty on VaR by a simulation study. Three scenarios are considered: noise in $\hat{\rho}$, noise in

$\widehat{PD}$ and noise in both. The results shows that the uncertainty of $\hat{\rho}$ generally brings in positive add-on, and this add-on gets significantly high when 1) the estimate of PD is high and the asset correlation is expected to be low or 2) the PD estimate is low but $\hat{\rho}$ is high. These trends get more obvious as the uncertainty of $\hat{\rho}$ increases. Moreover, the uncertainty of $\widehat{PD}$ brings only very small add-on. If the uncertainty of both are taken into account, the changes of add-on is similar to the finding in first scenario, with some difference that (a) the add-on is averagely larger and (b) the correct VaR suggests significantly higher economic capital when $\hat{\rho}$ is large, since the uncertainty of $\widehat{PD}$ is also large in this case, due to our simulation assumption.

Appendix

Appendix A:

Assume that

$$L_N|M \sim N(E(L_N|M), v^2)$$

where $M \sim N(0,1)$ and $v^2 \to 0$ as $N \to \infty$. The CDF of the unconditional loss can be written as

$$P(L_N \leq y) = E[P(L_N \leq y|M)]$$

where, for certain $\varepsilon > 0$,

$$P(L_N \leq y|M) = \begin{cases} 0, & y < E(L_N|M) - \varepsilon \\ 1, & y \geq E(L_N|M) + \varepsilon \end{cases}$$

This $\varepsilon \to 0$ when $v^2 \to 0$. Therefore,

$$P(L_N \leq y) = 1 \cdot P\big(y \geq E(L_N|M)\big), N \to \infty$$

We have know that

$$E(L_N|M) = \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot M}{\sqrt{1-\rho}}\right) = g(M)$$

Hence, asymptotically, $P(L_N \leq y) = P(E(L_N|M) \leq y) = P(g(M) \leq y) = P(M \geq g^{-1}(y))$. Now, set $P(L_N \leq y) = P(M \geq g^{-1}(y)) = q$ so that $y = VaR_q(L_N)$ and

$$g^{-1}(y) = \Phi^{-1}(1-q) \quad \leftrightarrow \quad y = g\big(\Phi^{-1}(1-q)\big).$$

This indicates that

$$VaR_q\big(E(L_N|M)\big) = VaR_q(L_N) = g\big(\Phi^{-1}(1-q)\big) = \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot \Phi^{-1}(1-q)}{\sqrt{1-\rho}}\right).$$

Appendix B: Relationship between the Default Correlation and the Asset Correlation

Denote $V_{it} = \sqrt{\rho} \cdot M_t + \sqrt{1-\rho} \cdot \varepsilon_{it}$, $I_{it} = 1_{\{V_{it} < \omega\}}$, $I_{it} \sim Bernoulli(1, PD)\ for\ \forall i, t$

Given *t*,

$$Cov(I_{it}, I_{jt}) = E(I_{it} I_{jt}) - E(I_{it})E(I_{jt})$$

The possibility of the occurrence of the event $I_{it} I_{jt}$ is $\Pr(V_{it} < \omega, V_{jt} < \omega)$ where

$$\begin{pmatrix} V_{it} \\ V_{jt} \end{pmatrix} \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Denote V_{it} as x_1 and V_{jt} as x_2 without losing generality, the density function of the joint normal distribution for these two variables can be written as

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)}\right\}$$

and then

$$\begin{aligned} F(x_1 = \omega, x_2 = \omega) &= \int_{-\infty}^{\omega} \int_{-\infty}^{\omega} f(x_1, x_2)\, d\, x_1 dx_2 \\ &= \int_{-\infty}^{\omega} A \left[\int_{-\infty}^{\omega} \exp\left\{ \frac{-x_1^2 + 2\rho x_1 x_2}{2(1-\rho^2)} \right\} dx_1 \right] dx_2 \\ &= \int_{-\infty}^{\omega} A \left[\int_{-\infty}^{\omega} \exp\{-\beta_1 x_1^2 + \beta_2 x_1 x_2\}\, dx_1 \right] dx_2 \\ &= \int_{-\infty}^{\omega} AB\ \ dx_2 \end{aligned}$$

where

$$A = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{-x_2^2}{2(1-\rho^2)}\right\} = a \cdot \exp\left\{\frac{-x_2^2}{2(1-\rho^2)}\right\}$$

$$\beta_1 = \frac{1}{2(1-\rho^2)}, \beta_2 = \frac{\rho^2}{1-\rho^2}$$

Now only consider about solving *B*,

$$B = \int_{-\infty}^{\omega} \exp\{-\beta_1 x_1^2 + \beta_2 x_1 x_2\}\, dx_1$$

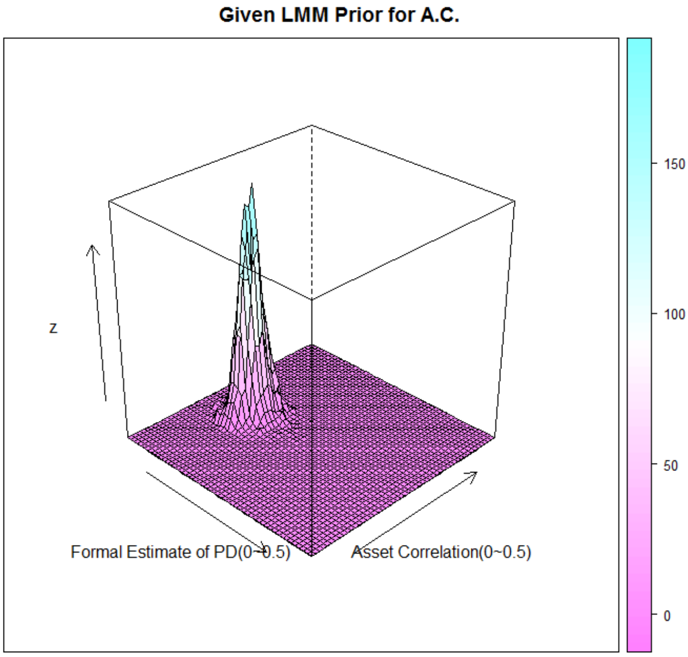
$$= \int_{-\infty}^{\omega} \exp\{-\beta_1 x_1^2\}\, dx_1 + \int_{-\infty}^{\omega} \exp\{\beta_2 x_1 x_2\}\, dx_1 = B_1 + B_2$$

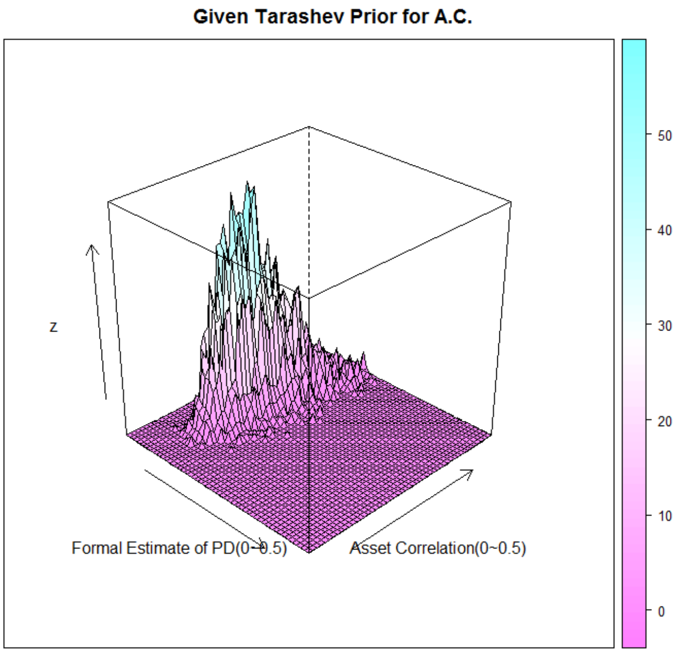
where B_2 can be easily solved but B_1 is hard to derive in an analytical form. However, it is easy to use a simulation for deriving the relationship between the default correlation and the asset correlation.

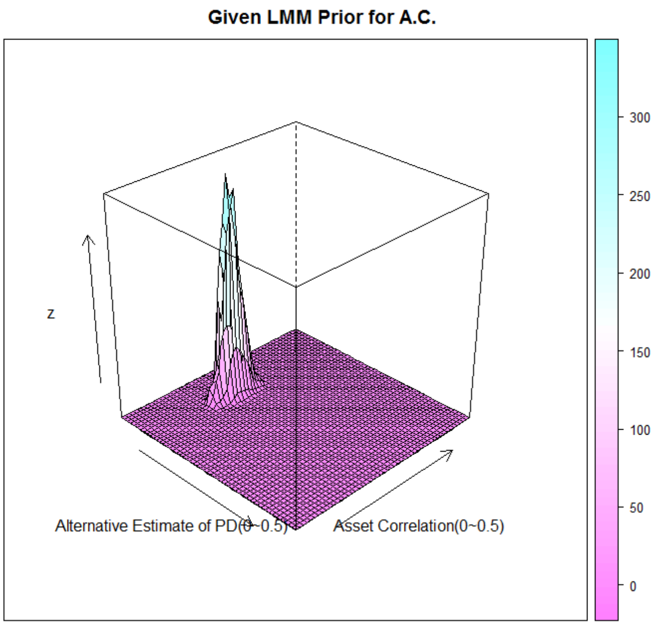
Default Threshold (ω)	Asset Correlation	Default Correlation
$\Phi^{-1}(0.05)$	0.10	0.0267
	0.20	0.0551
	0.30	0.0950
	0.40	0.1479
$\Phi^{-1}(0.10)$	0.10	0.0357
	0.20	0.0787
	0.30	0.1290
	0.40	0.1841
$\Phi^{-1}(0.20)$	0.10	0.0492
	0.20	0.1056
	0.30	0.1638
	0.40	0.2278

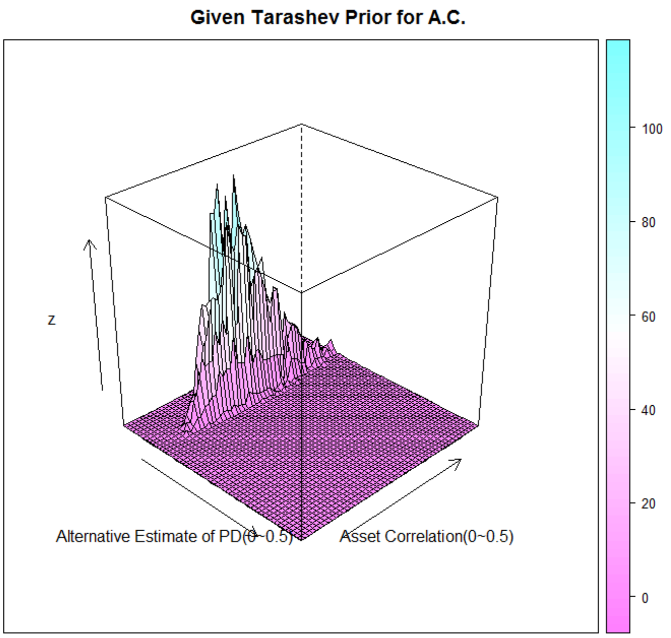
* For calculating each value of the Default correlation, first simulate 1000 pairs of default vectors (of dimension 1×500), and then calculate the correlation for each pair and take the mean.

Appendix C:


Given LMM Prior for A.C.
z
Formal Estimate of PD(0~0.5)
Asset Correlation(0~0.5)
150
100
50
0


Given Tarashev Prior for A.C.
z
Formal Estimate of PD(0~0.5)
Asset Correlation(0~0.5)
50
40
30
20
10
0


Given LMM Prior for A.C.
z
Alternative Estimate of PD(0~0.5)
Asset Correlation(0~0.5)
300
250
200
150
100
50
0


Given Tarashev Prior for A.C.
z
Alternative Estimate of PD(0~0.5)
Asset Correlation(0~0.5)
100
80
60
40
20
0

Appendix D:

Table 16 (Part 1). VaR Add-on,
Consider the Uncertainty of $\hat{\rho}$ only

$sd(\rho)$	$E(\rho)$	PD 0.05	0.1	0.15	0.2	0.25
0.1	0.1	0.022 (0.019,0.026)	0.04 (0.036,0.044)	0.054 (0.049,0.058)	0.063 (0.057,0.068)	0.069 (0.064,0.074)
	0.2	0.016 (0.012,0.019)	0.03 (0.025,0.034)	0.037 (0.031,0.045)	0.044 (0.039,0.048)	0.046 (0.04,0.051)
	0.3	0.013 (0.007,0.019)	0.026 (0.019,0.032)	0.033 (0.027,0.039)	0.034 (0.028,0.04)	0.035 (0.031,0.038)
	0.4	0.016 (0.01,0.023)	0.026 (0.019,0.033)	0.032 (0.024,0.039)	0.032 (0.027,0.036)	0.03 (0.025,0.034)
	0.5	0.018 (0.01,0.026)	0.029 (0.02,0.039)	0.032 (0.027,0.037)	0.028 (0.024,0.032)	0.025 (0.022,0.028)
	0.6	0.026 (0.016,0.035)	0.035 (0.029,0.04)	0.032 (0.027,0.036)	0.025 (0.022,0.027)	0.018 (0.017,0.019)
0.15	0.1	0.033 (0.028,0.037)	0.064 (0.057,0.07)	0.089 (0.081,0.095)	0.105 (0.097,0.115)	0.116 (0.111,0.122)
	0.2	0.03 (0.025,0.036)	0.058 (0.051,0.064)	0.075 (0.067,0.082)	0.084 (0.077,0.091)	0.088 (0.082,0.094)
	0.3	0.03 (0.023,0.035)	0.056 (0.048,0.064)	0.067 (0.06,0.074)	0.074 (0.068,0.078)	0.073 (0.068,0.077)
	0.4	0.032 (0.023,0.038)	0.058 (0.048,0.069)	0.067 (0.059,0.075)	0.067 (0.063,0.073)	0.061 (0.056,0.065)
	0.5	0.04 (0.031,0.051)	0.065 (0.057,0.074)	0.068 (0.062,0.074)	0.061 (0.057,0.065)	0.049 (0.046,0.051)
	0.6	0.058 (0.046,0.073)	0.076 (0.07,0.084)	0.064 (0.06,0.069)	0.046 (0.044,0.048)	0.031 (0.03,0.032)

Table 16 (Part 2). VaR Add-on,
Consider the Uncertainty of $\hat{\rho}$ only

$sd(\rho)$	$E(\rho)$	PD 0.05	0.1	0.15	0.2	0.25
0.2	0.1	0.038 (0.032,0.044)	0.084 (0.076,0.091)	0.12 (0.111,0.129)	0.147 (0.138,0.155)	0.164 (0.155,0.173)
	0.2	0.046 (0.04,0.05)	0.089 (0.082,0.097)	0.118 (0.111,0.126)	0.133 (0.122,0.139)	0.139 (0.131,0.145)
	0.3	0.048 (0.04,0.055)	0.09 (0.082,0.097)	0.114 (0.105,0.124)	0.119 (0.111,0.126)	0.115 (0.11,0.119)
	0.4	0.056 (0.045,0.067)	0.1 (0.091,0.109)	0.113 (0.105,0.12)	0.108 (0.102,0.113)	0.096 (0.092,0.1)
	0.5	0.072 (0.057,0.086)	0.114 (0.103,0.124)	0.112 (0.106,0.117)	0.092 (0.089,0.095)	0.07 (0.069,0.072)
	0.6	0.101 (0.089,0.117)	0.128 (0.119,0.138)	0.095 (0.093,0.098)	0.061 (0.06,0.061)	0.037 (0.036,0.037)
0.25	0.1	0.038 (0.031,0.044)	0.098 (0.089,0.108)	0.148 (0.138,0.158)	0.19 (0.176,0.198)	0.214 (0.204,0.226)
	0.2	0.059 (0.052,0.068)	0.123 (0.112,0.133)	0.164 (0.15,0.175)	0.186 (0.178,0.197)	0.19 (0.182,0.197)
	0.3	0.069 (0.057,0.081)	0.134 (0.126,0.144)	0.162 (0.154,0.168)	0.169 (0.161,0.177)	0.159 (0.153,0.165)
	0.4	0.082 (0.073,0.093)	0.148 (0.137,0.158)	0.164 (0.155,0.171)	0.15 (0.145,0.156)	0.126 (0.124,0.129)
	0.5	0.112 (0.099,0.127)	0.17 (0.16,0.18)	0.154 (0.15,0.159)	0.117 (0.115,0.118)	0.082 (0.082,0.083)
	0.6	0.164 (0.143,0.181)	0.175 (0.17,0.18)	0.111 (0.11,0.111)	0.064 (0.064,0.064)	0.037 (0.037,0.037)
0.3	0.1	0.033 (0.028,0.039)	0.109 (0.1,0.121)	0.181 (0.167,0.195)	0.232 (0.217,0.247)	0.269 (0.261,0.277)
	0.2	0.07 (0.06,0.078)	0.16 (0.148,0.171)	0.216 (0.206,0.228)	0.24 (0.229,0.251)	0.243 (0.234,0.252)
	0.3	0.088 (0.077,0.102)	0.182 (0.17,0.193)	0.221 (0.213,0.229)	0.221 (0.212,0.229)	0.199 (0.195,0.202)
	0.4	0.115 (0.101,0.132)	0.207 (0.194,0.22)	0.215 (0.207,0.224)	0.184 (0.181,0.188)	0.145 (0.144,0.147)
	0.5	0.16 (0.144,0.178)	0.228 (0.221,0.236)	0.183 (0.179,0.186)	0.127 (0.126,0.127)	0.085 (0.085,0.085)
	0.6	0.235 (0.219,0.256)	0.2 (0.198,0.202)	0.113 (0.113,0.113)	0.064 (0.064,0.064)	0.037 (0.037,0.037)

Table 17. VaR Add-on,
Consider the Uncertainty of $\widehat{PD}$ only

ρ	E(PD) 0.05	0.1	0.15	0.2	0.25
0.1	-0.004 (-0.084,0.051)	0.000 (-0.08,0.058)	-0.004 (-0.052,0.035)	-0.003 (-0.048,0.029)	-0.003 (-0.043,0.024)
0.2	-0.012 (-0.178,0.139)	-0.004 (-0.148,0.112)	-0.013 (-0.091,0.045)	0.007 (-0.047,0.067)	-0.001 (-0.066,0.043)
0.3	0.004 (-0.228,0.188)	-0.013 (-0.134,0.115)	-0.025 (-0.135,0.08)	-0.007 (-0.084,0.071)	-0.001 (-0.071,0.057)
0.4	0.009 (-0.285,0.228)	-0.002 (-0.236,0.156)	-0.022 (-0.18,0.082)	-0.02 (-0.144,0.054)	-0.012 (-0.055,0.047)
0.5	-0.012 (-0.348,0.228)	-0.006 (-0.221,0.138)	-0.03 (-0.246,0.099)	-0.008 (-0.089,0.053)	-0.003 (-0.064,0.045)
0.6	-0.004 (-0.443,0.304)	-0.055 (-0.362,0.124)	-0.024 (-0.165,0.069)	-0.019 (-0.121,0.037)	-0.006 (-0.092,0.025)

Table 18 (Part 1). VaR Add-on,
Consider the Uncertainty of both $\hat{\rho}$ and $\widehat{PD}$

$sd(\rho)$	$E(\rho)$	$E(PD)$ 0.05	0.1	0.15	0.2	0.25
0.1	0.1	0.043 (0.033,0.051)	0.051 (0.037,0.063)	0.06 (0.046,0.076)	0.065 (0.047,0.082)	0.075 (0.054,0.092)
	0.2	0.052 (0.044,0.062)	0.046 (0.036,0.055)	0.05 (0.038,0.065)	0.05 (0.036,0.065)	0.052 (0.038,0.066)
	0.3	0.069 (0.06,0.078)	0.05 (0.041,0.063)	0.046 (0.035,0.057)	0.045 (0.034,0.056)	0.043 (0.034,0.052)
	0.4	0.095 (0.087,0.107)	0.055 (0.045,0.064)	0.048 (0.035,0.059)	0.042 (0.033,0.051)	0.036 (0.03,0.044)
	0.5	0.125 (0.113,0.137)	0.064 (0.053,0.076)	0.046 (0.038,0.056)	0.037 (0.03,0.044)	0.03 (0.022,0.036)
	0.6	0.159 (0.144,0.174)	0.07 (0.061,0.081)	0.044 (0.038,0.052)	0.031 (0.026,0.035)	0.02 (0.017,0.023)
0.15	0.1	0.06 (0.048,0.077)	0.078 (0.056,0.101)	0.092 (0.066,0.12)	0.112 (0.083,0.135)	0.126 (0.09,0.158)
	0.2	0.072 (0.056,0.087)	0.078 (0.058,0.101)	0.087 (0.066,0.104)	0.096 (0.079,0.112)	0.094 (0.078,0.121)
	0.3	0.094 (0.079,0.116)	0.08 (0.061,0.097)	0.084 (0.065,0.104)	0.085 (0.069,0.099)	0.079 (0.063,0.094)
	0.4	0.122 (0.102,0.137)	0.09 (0.069,0.109)	0.081 (0.064,0.093)	0.073 (0.059,0.085)	0.067 (0.057,0.079)
	0.5	0.16 (0.134,0.182)	0.102 (0.085,0.12)	0.084 (0.069,0.099)	0.067 (0.058,0.075)	0.053 (0.047,0.059)
	0.6	0.206 (0.188,0.226)	0.11 (0.094,0.125)	0.075 (0.067,0.081)	0.049 (0.043,0.054)	0.032 (0.03,0.034)

Table 18 (Part 2). VaR Add-on,
Consider the Uncertainty of both $\hat{\rho}$ and $\widehat{PD}$

$sd(\rho)$	$E(\rho)$	$E(PD)$ 0.05	0.1	0.15	0.2	0.25
0.2	0.1	0.069 (0.046,0.091)	0.1 (0.066,0.132)	0.131 (0.093,0.181)	0.162 (0.119,0.196)	0.173 (0.13,0.221)
	0.2	0.098 (0.075,0.119)	0.116 (0.093,0.142)	0.129 (0.102,0.157)	0.138 (0.11,0.164)	0.14 (0.11,0.178)
	0.3	0.124 (0.102,0.149)	0.118 (0.1,0.137)	0.128 (0.107,0.154)	0.127 (0.105,0.148)	0.122 (0.102,0.144)
	0.4	0.16 (0.133,0.189)	0.136 (0.104,0.161)	0.129 (0.103,0.15)	0.119 (0.103,0.132)	0.101 (0.088,0.112)
	0.5	0.21 (0.17,0.244)	0.149 (0.128,0.172)	0.125 (0.109,0.143)	0.097 (0.084,0.107)	0.072 (0.066,0.078)
	0.6	0.259 (0.239,0.283)	0.155 (0.14,0.169)	0.1 (0.094,0.106)	0.062 (0.059,0.063)	0.037 (0.036,0.037)
0.25	0.1	0.071 (0.052,0.099)	0.12 (0.083,0.167)	0.165 (0.115,0.206)	0.2 (0.147,0.248)	0.23 (0.163,0.277)
	0.2	0.114 (0.089,0.146)	0.154 (0.113,0.197)	0.185 (0.145,0.216)	0.19 (0.161,0.231)	0.194 (0.157,0.227)
	0.3	0.157 (0.125,0.193)	0.17 (0.145,0.198)	0.182 (0.152,0.212)	0.179 (0.152,0.207)	0.165 (0.145,0.185)
	0.4	0.212 (0.174,0.257)	0.192 (0.166,0.221)	0.178 (0.15,0.197)	0.156 (0.14,0.173)	0.127 (0.116,0.137)
	0.5	0.268 (0.235,0.3)	0.208 (0.179,0.235)	0.162 (0.152,0.172)	0.118 (0.11,0.123)	0.083 (0.08,0.084)
	0.6	0.321 (0.299,0.349)	0.19 (0.178,0.198)	0.112 (0.11,0.113)	0.064 (0.064,0.064)	0.037 (0.037,0.037)
0.3	0.1	0.076 (0.046,0.11)	0.137 (0.087,0.199)	0.194 (0.137,0.248)	0.238 (0.164,0.303)	0.275 (0.204,0.352)
	0.2	0.139 (0.101,0.177)	0.191 (0.15,0.241)	0.233 (0.188,0.292)	0.247 (0.203,0.293)	0.248 (0.209,0.281)
	0.3	0.195 (0.148,0.241)	0.219 (0.19,0.252)	0.239 (0.205,0.277)	0.221 (0.181,0.252)	0.202 (0.188,0.217)
	0.4	0.258 (0.214,0.299)	0.252 (0.218,0.283)	0.226 (0.201,0.25)	0.187 (0.176,0.198)	0.145 (0.139,0.15)
	0.5	0.33 (0.295,0.373)	0.256 (0.238,0.276)	0.186 (0.176,0.191)	0.127 (0.125,0.128)	0.085 (0.084,0.085)
	0.6	0.37 (0.352,0.383)	0.204 (0.201,0.205)	0.113 (0.113,0.113)	0.064 (0.064,0.064)	0.037 (0.037,0.037)

Reference

Crouhy, M., Galai, D., Mark, R. (2000). A comparative analysis of current credit risk models. Journal of Banking&Finance, issue 24: 59-117.

Dwyer, D. W. (2007). The distribution of defaults and Bayesian model validation. Journal of Risk Model Validation, vol. 1(1): 23-53.

Etheridge, A. (2002). A course in financial calculus. Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H., Rubin, D. B. (2003). Bayesian data analysis, second edition. Chapman and Hall/CRC.

Gelman, A., Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Gilks, W. R., Richardson, S., Spiegelhalter, D. (1995). Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC.

Gordy, M. B., Lutkebohmert, E. (2007). Granularity adjustment for Basel II. Deutsche Bundesbank Discussion Paper, No.01.

Gordy, M. B. (2000). A comparative anatomy of credit risk models. Journal of Banking &Finance, issue 24(1-2): 119-149.

Gordy, M. B. (2003). A risk-factor model foundation for rating-based bank capital rules. Journal of Financial Intermediation, vol. 12: 199-232.

Hardle, W. K., Hautsch, N., Overbeck, L. (2008). Applied quantitative finance, second edition. Springer Berlin Heidelberg.

Heitfield, E. (2008). Parameter uncertainty and the credit risk of collateralized debt obligations. Working paper, Federal Reserve Board.

Kupiec, P. (2008). A generalized single common factor model of portfolio credit risk. Journal of Derivatives, vol 15: 25-40.

McCulloch, C. E., Searle, S. R., Neuhaus, J. M. (2008). Generalized, linear and mixed models, second edition. Wiley-Blackwell press.

McCullagh, P., Nelder, J. A. (1989). Generalized linear models, second edition. Chapman and Hall/CRC.

McNeil, A. J., Wendin, J. (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. Journal of Empirical Finance, vol. 14 (2): 131-149.

Saita, F. (2007). Value at risk and bank capital management. Elsevier Inc.

Tarashev, N., Zhu, H. (2008). Specification and calibration errors in measures of portfolio credit risk: the case of the ASRF model. International Journal of Central Banking, vol. 4 (2): 129-173.

Tarashev, N. (2010). Measuring portfolio credit risk correctly: why parameter uncertainty matters. Journal of Banking & Finance, issue 34: 2065-2076.

Vasicek, O. (1991). Limiting loan loss probability distribution, KMV Working Paper.