Analysis of Credit Portfolio Risk using Hierarchical Multi-Factor Models

Pak-Wing Fok,
Department of Mathematical Sciences,
University of Delaware,
Newark,
DE 19711
pakwing@udel.edu

Xiuling Yan,
Department of Accounting and Management Information Systems,
University of Delaware,
Newark,
DE 19711
xyan@udel.edu

Guangming Yao,
Department of Mathematics,
Clarkson University,
Potsdam,
NY 13699
guangmingyao@gmail.com

September 2, 2014

Abstract

In this paper we generalize Vasicek's Asymptotic Single Risk Factor (ASRF) solution to multiple factors organized with a particular hierarchical structure. We use this model to investigate credit portfolio loss. In this hierarchical factor model, asset returns of a company depend on a global factor, a sector factor, and an idiosyncratic risk factor. All companies share the same global factor and all companies within a sector share the same sector factor. Using the central limit theorem, we derive closed form solutions for the Value-at-Risk (VaR) and expected shortfall (ES) under the assumptions that the number of sectors in the portfolio is large, and the exposures scale as the reciprocal of the number of sectors. Our results for the VaR agree well with Monte-Carlo simulations providing the sector factor loadings and variance of systematic risk are not too large.

Keywords: Vasicek Model; Asymptotic Single Risk Factor; Factor Models; Correlated Default; Credit Portfolio

1 Introduction

Credit portfolios are portfolios of fixed-income investment products such as bonds, loans, and credit derivatives. Fixed-income investment products provide the investor with a steady stream of cash inflow (e.g. in form of interest payments) during the lifetime of the product. The trade-off is limited upward potential for gain of the portfolio. Banks, insurance companies and other financial institutions regularly maintain and manage large numbers of credit portfolios. The main risk associated with such portfolios is when a debtor defaults on its obligation. Although such an event is rare, a single default often means that the entire portfolio goes to loss. Therefore, investors in credit portfolios need systematic methods to analyze the associated risk and to create financial instruments to insure against losses, should they arise. The management of credit risk is a vital area of research within quantitative finance; see for example Bohn and Stein (2009); Denault (2001); McNeil et al. (2010); Servigny and Renault (2004).

It is well-known that companies do not default independently from each other (Lucas, 1995). One common way of modeling correlated company defaults is through *factor models*, (Bluhm et al., 2010; Schönbucher, 2001; Burtschell et al., 2009). In these models, a representation for the asset return of a company is specified in terms of random variables. When the asset return drops below a given threshold, the company defaults. Correlation in company default is included by allowing the random variables to share common factors. Factor models are studied and commonly employed by companies such as Moody's KMV (Crouhy et al., 2000) and the RiskMetrics group (Gordy, 2003).

The Vasicek credit model (Vasicek, 1987) provides a simple analytic solution for a portfolio containing identical companies that are coupled to a single global factor. While Vasicek's Asymptotic Single Risk Factor (ASRF) solution is simple to derive, it also can be easily extended to the heterogeneous case and importantly, forms the foundation of the Basel accords for bank capital requirements (Basel Committee on Banking Supervision, 2006). Other authors have extended the ASRF model to account for uncertainty over loss given default (Kupiec, 2008) and multiple global risk factors (Schönbucher, 2001; Pykhtin, 2004).

In this paper, we propose and validate an analytic formula for the loss distribution of a credit portfolio, assuming a hierarchical multi-factor model. In such a model, all companies have exposure to a global risk factor; in addition, companies in a given sector are subject to a local risk factor. Thus it may be regarded as a simple extension of the ASRF model to an economically intuitive multi-factor case. Since our loss formula can be written entirely in terms of elementary and special functions, it is much quicker to evaluate than Monte-Carlo simulations which are often time consuming and computationally expensive. Our derivation involves analyzing the sector loss and then applying the central limit theorem to all the sectors. It is similar to saddle point methods (Huang et al., 2007; Jensen, 1995; Lugannani

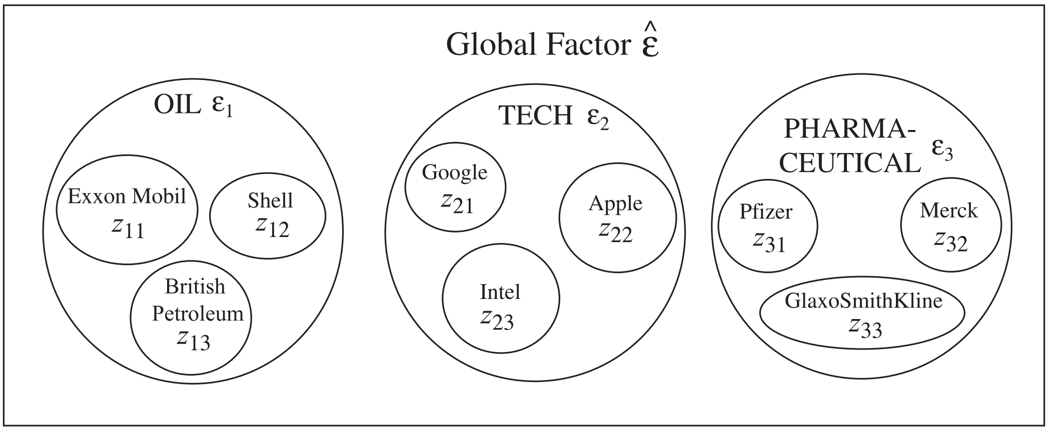

Figure 1: Hierarchical factor model of company asset returns within a global economy with $N = 3$ sectors and $n = 3$ companies per sector. Every company participates in a global economy and belongs to exactly one sector. Each company's asset return z_{ij} is affected by a global factor $\hat{\varepsilon}$, a sector factor ε_i and an idiosyncratic factor ζ_{ij} (not shown). See eq. (1).

and Rice, 1980) in the sense that both methods approximate sums of random variables through asymptotic formulas.

The layout of this paper is as follows. In section 2, we introduce the hierarchical factor model for a firm's value and set up the portfolio in terms of individual companies and their default probabilities. In section 3, we derive the value-at-risk (VaR) for a portfolio coupled to a hierarchical factor model. In section 4 we compare our solutions for the VaR with Monte Carlo simulations. We conclude the paper in section 5.

2 Portfolio Pricing in a Localized 1-Factor Model

One critical issue that determines the value of credit portfolios is the default correlation among companies (Schönbucher, 2001). Although the default probability of a company may be very small, defaults between companies are often correlated. Factor models incorporate the correlation among asset returns explicitly by assuming that they are driven by M shared "factors" which are modeled as independent random variables. For example, in an $M = 2$ factor model, these shared factors could represent the state of a country's economy (a recession negatively impacts all companies in that country), or the price of a resource (a lower price would lower the expenses of all companies that use the resource). All asset returns in an $M = 2$ factor model would be influenced by the country's economy and the price of the resource.

In this paper we restrict our attention to hierarchical (or 'localized') factor models which have the advantage of being simple, yet economically intuitive. In such models, the asset return for a company depends on a *global* factor that is shared by all companies and exactly one of N other *sector* factors: see Fig. 1. Hence, all companies are correlated through the global factor and all companies in the same sector are further correlated. Although in this paper our model partitions a portfolio into different industrial sectors, our approach can also be applied to partitions of geography and size buckets.

Specifically, using a Merton model for a firm's value (Merton, 1974), we consider a special case of a $N+1$ factor model where the normalized asset return of the jth company in sector i is given by

$$z_{ij} = \sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon} + \beta_{ij}\varepsilon_i) + \sqrt{1-\rho_{ij}}\zeta_{ij}, \qquad i = 1,\ldots,N, \quad j = 1,\ldots,n, \tag{1}$$

and $0 < \rho_{ij} < 1$, $0 < \beta_{ij}, \hat{\beta}_{ij} < 1$. In this model, the asset returns are normalized so that $\mathrm{Var}(z_{ij}) = 1$, i.e. $\hat{\beta}_{ij}^2 + \beta_{ij}^2 = 1$. The $\hat{\beta}_{ij}$ and β_{ij} are the global and sector *factor loadings*; knowledge of β_{ij} determines $\hat{\beta}_{ij}$ and vice-versa. Since $\sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon}+\beta_{ij}\varepsilon_i)$ is the systematic risk, ρ_{ij} is the *variance of systematic risk*, $\hat{\varepsilon} \sim \mathcal{N}(0,1)$ is the *global risk factor*, $\{\varepsilon_i\} \sim \mathcal{N}(0,1)$ are the N independent *sector risk factors* and $\{\zeta_{ij}\} \sim \mathcal{N}(0,1)$ are the $N \times n$ independent *idiosyncratic risk factors*.

Although there are $N+1$ factors altogether, each company only depends on two of them. Any two companies are always correlated at the global level, and possibly also at the sector level. Specifically, we have

$$\mathrm{Corr}(z_{ij}, z_{kl}) = \begin{cases} \sqrt{\rho_{ij}\rho_{kl}}\hat{\beta}_{ij}\hat{\beta}_{kl}, & \text{if } i \neq k, \\ \sqrt{\rho_{ij}\rho_{il}}(\hat{\beta}_{ij}\hat{\beta}_{il} + \beta_{ij}\beta_{il}), & \text{if } i = k. \end{cases} \tag{2}$$

In other words, the asset returns of any two companies are always correlated through their global factor loadings. If the companies also happen to be in the same sector, they are further correlated through their sector factors.

We point out that the hierarchical factor model for the company's asset return (1) is a special case of a multi-factor model where all companies are influenced (in various degrees) by multiple systematic risk factors. This general case has been studied by Pykhtin (2004); his approach is to optimally approximate the multifactor model by carefully choosing the factor loadings in the single factor model. In our approach, we assume a simpler structure for the company's asset return from the very beginning. In return, we obtain an analytic solution which is simple to evaluate, and is exact in the limit as $N \to \infty$.

We assume that a company defaults if its asset return z_{ij} drops below a threshold value θ_{ij}. In principle all companies within the global economy could have different thresholds and details of how to determine them can be found in Crosbie and Bohn (1999). Over a fixed time horizon, we may write the loss on the portfolio as

$$R_\pi = \sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij} R_{ij}(z_{ij}), \tag{3}$$

where the exposures $w_{ij} > 0$ satisfy $\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij} = 1$. More dynamic evolution models that explicitly account for the Brownian nature of a firm's value are studied in Huh and Kolkiewicz (2008); Hurd (2009); Li (2000). We model the portfolio loss as a mixture of Bernoulli random variables (Joe, 1997) by assuming a constant percentage loss given default (LGD):

$$R_{ij}(z_{ij}) = \begin{cases} 0, & \text{if } z_{ij} > \theta_{ij}, \\ c, & \text{if } z_{ij} \leq \theta_{ij}, \end{cases} \tag{4}$$

where $0 < c \leq 1$. As a consequence, $0 \leq R_\pi \leq c$. Because the idiosyncratic factor is unit-normally distributed, the conditional probability of default for company j in sector i is

$$\text{Prob}(z_{ij} < \theta_{ij} |\ \hat{\varepsilon}, \varepsilon_i) = p_{ij}(\varepsilon_i, \hat{\varepsilon}) = \Phi\left[\frac{\theta_{ij} - \sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon} + \beta_{ij}\varepsilon_i)}{\sqrt{1-\rho_{ij}}}\right], \tag{5}$$

where $\Phi(z) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{z}{\sqrt{2}}\right)\right]$. Throughout this paper, we use $\phi(z) = \Phi'(z) = \exp[-z^2/2]/\sqrt{2\pi}$. As a simple corollary to (5), since $\int_{-\infty}^{\infty} \Phi(a + b\varepsilon)\phi(\varepsilon)\text{d}\varepsilon = \Phi(a/\sqrt{1+b^2})$ for constants a and b, the (unconditional) default probability of the jth company in sector i is

$$\begin{aligned} PD_{ij} = \text{Prob}(z_{ij} < \theta_{ij}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi\left[\frac{\theta_{ij} - \sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon} + \beta_{ij}\varepsilon_i)}{\sqrt{1-\rho_{ij}}}\right] \phi(\hat{\varepsilon})\phi(\varepsilon_i)\text{d}\hat{\varepsilon}\text{d}\hat{\varepsilon}_i, \\ &= \Phi(\theta_{ij}). \end{aligned} \tag{6}$$

Hence, the default probability is uniquely determined by specifying the threshold value θ_{ij}.

Implementation of the factor model now requires knowledge of the numerical values of β_{ij}, $\hat{\beta}_{ij}$, ρ_{ij} and θ_{ij}. The exposures w_{ij} can be chosen by the user of the model, or taken from the call reports of banks which can be found from the website of the Federal Deposit Insurance Corporation (FDIC). The θ_{ij} are related to the default probabilty through eq. (6) and the default probabilities can be inferred from the credit rating of a firm (Crouhy et al., 2000). The correlation terms $\hat{\beta}_{ij}$, $\hat{\beta}_{ij}$ and ρ_{ij} are more difficult to obtain, but correlation matrices can usually be found empirically through historical data. See Andersen et al. (2003) for more details.

3 Loss Distribution for the Portfolio

The main result in this section is a derivation for the Value-at-Risk (VaR) for R_π in (3). In our analysis we assume that the number of sectors is large: $N \gg 1$ and the positive exposures $w_{ij} = O(N^{-1})$ as $N \to \infty$. Although the $N \gg 1$ assumption may be somewhat unrealistic, the value of N does depend on how the loans are grouped together. If historical data and economic intuition can allow a different grouping with a larger N, this would be lead to a more accurate model, according to the analysis in this paper. The number of companies per sector n is typically very large (perhaps $n \gtrsim 1000$), but this is not a necessary requirement for our approximations to hold.

We first provide some preliminary results for the moments of R_π before proving the main theorem.

Lemma 1. *Let $R_\pi = \sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}R_{ij}(z_{ij})$ where $R_{ij}(\cdot)$ satisfies eq. (4). Furthermore, let $Y_i = \sum_{j=1}^{n} w_{ij}R_{ij}(z_{ij})$, $i = 1, \ldots, N$ be the sector losses so that $R_\pi = \sum_{i=1}^{N} Y_i$ and let the* *exposures* $w_{ij} = O(N^{-1})$ *as* $N \to \infty$*. Define the conditional moments*

$$\begin{aligned}
\mu_i(\hat{\varepsilon}, \varepsilon_i) &= E[Y_i|\hat{\varepsilon}, \varepsilon_i], &(7)\\
\sigma_i^2(\hat{\varepsilon}, \varepsilon_i) &= V[Y_i|\hat{\varepsilon}, \varepsilon_i], &(8)\\
m_i(\hat{\varepsilon}) &= E[Y_i|\hat{\varepsilon}], &(9)\\
s_i^2(\hat{\varepsilon}) &= V[Y_i|\hat{\varepsilon}], &(10)\\
m(\hat{\varepsilon}) &= E[R_\pi|\hat{\varepsilon}], &(11)\\
s^2(\hat{\varepsilon}) &= V[R_\pi|\hat{\varepsilon}]. &(12)
\end{aligned}$$

Then, as $N \to \infty$*, the conditional means satisfy*

$$\begin{aligned}
\mu_i(\hat{\varepsilon}, \varepsilon_i) &= c\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) - \sqrt{\rho_{ij}}(\hat{\varepsilon}\hat{\beta}_{ij} + \varepsilon_i\beta_{ij})}{\sqrt{1-\rho_{ij}}}\right] = O(N^{-1}), &(13)\\
m_i(\hat{\varepsilon}) &= c\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) - \sqrt{\rho_{ij}}\hat{\varepsilon}\hat{\beta}_{ij}}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right] = O(N^{-1}), &(14)\\
m(\hat{\varepsilon}) &= c\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) - \sqrt{\rho_{ij}}\hat{\varepsilon}\hat{\beta}_{ij}}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right] = O(1), &(15)
\end{aligned}$$

while the conditional variances scale as

$$\begin{aligned}
\sigma_i^2(\varepsilon_i, \hat{\varepsilon}) &= O(N^{-2}), &(16)\\
s_i^2(\hat{\varepsilon}) &= O(N^{-2}), &(17)\\
s^2(\hat{\varepsilon}) &= O(N^{-1}). &(18)
\end{aligned}$$

Proof. Conditioned on $\hat{\varepsilon}$ and ε_i, the $R_{ij}(z_{ij})$ are independent (scaled) Bernoulli random variables, so eq. (13) follows from

$$\begin{aligned}
\mu_i(\hat{\varepsilon}, \varepsilon_i) &= \sum_{j=1}^{n} w_{ij}E[R_{ij}(z_{ij})] = c\sum_{j=1}^{n} w_{ij}p_{ij},\\
&= c\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) - \sqrt{\rho_{ij}}(\hat{\varepsilon}\hat{\beta}_{ij} + \varepsilon_i\beta_{ij})}{\sqrt{1-\rho_{ij}}}\right] = O(N^{-1}),
\end{aligned}$$

using eqs. (5) and (6). Eq. (14) follows from $m_i(\hat{\varepsilon}) = \int_{-\infty}^{\infty} \mu_i(\hat{\varepsilon}, \varepsilon_i)\phi(\varepsilon_i)\mathrm{d}\varepsilon_i$ and eq. (15) follows from $m(\hat{\varepsilon}) = \sum_{i=1}^{N} m_i(\hat{\varepsilon})$. Now we prove eqs. (16)-(18). Conditioned on $\hat{\varepsilon}$ and ε_i, the $R_{ij}(z_{ij})$ are again independent (scaled) Bernoulli random variables. With the shorthand $p_{ij}(\hat{\varepsilon}, \varepsilon_i) = \Phi\left[\frac{\theta_{ij} - \sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon} + \beta_{ij}\varepsilon_i)}{\sqrt{1-\rho_{ij}}}\right]$, eq. (16) follows from $\sigma_i^2(\hat{\varepsilon}, \varepsilon_i) = c^2\sum_{j=1}^{n} w_{ij}^2 p_{ij}(\hat{\varepsilon}, \varepsilon_i)[1 -$

$p_{ij}(\hat{\varepsilon},\varepsilon_i)] = O(N^{-2})$, eq. (17) follows from

$$\begin{aligned}
s_i^2(\hat{\varepsilon}) &= E[Y_i^2|\hat{\varepsilon}] - E[Y_i|\hat{\varepsilon}]^2, \\
&= \int_{-\infty}^{\infty} E[Y_i^2|\hat{\varepsilon},\varepsilon_i]\phi(\varepsilon_i)\mathrm{d}\varepsilon_i - \left(\int_{-\infty}^{\infty} E[Y_i|\hat{\varepsilon},\varepsilon_i]\phi(\varepsilon_i)\mathrm{d}\varepsilon_i\right)^2, \\
&= c^2\sum_{k=1}^{n}\sum_{j=1}^{n} w_{ij}w_{ik}\int_{-\infty}^{\infty} p_{ij}(\hat{\varepsilon},\varepsilon_i)p_{ik}(\hat{\varepsilon},\varepsilon_i)\phi(\varepsilon_i)\mathrm{d}\varepsilon_i - \left\{c\sum_{j=1}^{n} w_{ij}\int_{-\infty}^{\infty} p_{ij}(\hat{\varepsilon},\varepsilon_i)\phi(\varepsilon_i)\mathrm{d}\varepsilon_i\right\}^2, \\
&= O(N^{-2}),
\end{aligned}$$

and therefore eq. (18) immediately follows since $s^2(\hat{\varepsilon}) = \sum_{i=1}^{N} s_i^2(\hat{\varepsilon})$. □

Theorem 1. *Let the normalized asset return of the jth company in sector i follow a hierarchical multi-factor model so that the correlated random variables z_{ij} satisfy*

$$z_{ij} = \sqrt{\rho_{ij}}(\hat{\beta}_{ij}\hat{\varepsilon} + \beta_{ij}\varepsilon_i) + \sqrt{1-\rho_{ij}}\zeta_{ij}, \qquad i = 1,\ldots,N, \quad j = 1,\ldots,n, \tag{19}$$

where $0 < \rho_{ij} < 1$, $0 < \beta_{ij}, \hat{\beta}_{ij} < 1$, $\beta_{ij}^2 + \hat{\beta}_{ij}^2 = 1$ and $\varepsilon_i, \hat{\varepsilon} \sim \mathcal{N}(0,1)$. Consider a portfolio

$$R_\pi = \sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}R_{ij}(z_{ij}), \tag{20}$$

with exposures w_{ij} such that $\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij} = 1$ and $w_{ij} = O(N^{-1})$ as $N \to \infty$, where $R_{ij}(z_{ij})$ follows eq. (4): i.e. for some $-\infty < \theta_{ij} < \infty$, the company whose asset return follows eq. (19) defaults when $z_{ij} < \theta_{ij}$, incurring a loss c. Then the Value-at-Risk (VaR_q) of the portfolio R_π at risk level $0 \le q \le 1$ satisfies the asymptotic relation

$$VaR_q \sim c\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) + \sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q)}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right]. \tag{21}$$

as $N \to \infty$, where

$$VaR_q = \inf\left\{x : q \le F_{R_\pi}(x)\right\}, \tag{22}$$

and $F_{R_\pi}(\cdot)$ is the cumulative density function of R_π.

Proof. We write the loss of the portfolio as

$$R_\pi = \sum_{i=1}^{N} Y_i, \qquad Y_i = \sum_{j=1}^{n} w_{ij}R_{ij}(z_{ij}), \tag{23}$$

so that Y_i is the total loss of all companies in sector i and R_π is the loss summed over all sectors. Conditioned on the global risk $\hat{\varepsilon}$, the Y_i, $i = 1,\ldots,N$ are independent random variables and from the central limit theorem, their sum follows a normal distribution when $N \gg 1$: $R_\pi \equiv \sum_{i=1}^{N} Y_i \sim N\left[m(\hat{\varepsilon}), s^2(\hat{\varepsilon})\right]$, where $m(\cdot)$ is given by eq. (15) and s^2 obeys the scaling (18) (its explicit form is not required to derive (21)). The density for the loss of the entire portfolio is given by the law of total probability:

$$f_{R_\pi}(L) \quad \sim \quad \sqrt{N}\int_{-\infty}^{\infty} \frac{\phi(\hat{\varepsilon})}{\sqrt{2\pi\hat{s}^2(\hat{\varepsilon})}} \exp\left\{\frac{-N[L-m(\hat{\varepsilon})]^2}{2\hat{s}^2(\hat{\varepsilon})}\right\} \mathrm{d}\hat{\varepsilon}, \tag{24}$$

as $N \to \infty$, where we have set $s^2 = \hat{s}^2/N$ in light of eq. (18).

We now apply Laplace's method to the integral in (24). Laplace's method provides a way to approximate integrals that contain a large parameter by analyzing the *stationary point* of the integrand, and an overview of the method is given in the appendix. In particular we refer to eq. (44) which approximates integrals of the form $\int_a^b g(t)\exp[-k\psi(t)]\mathrm{d}t$ for functions g and ψ, with $a, b \in \mathbb{R}$ and $k \gg 1$. Taking $\psi(\hat{\varepsilon}) = (L - m(\hat{\varepsilon}))^2/(2s^2(\hat{\varepsilon}))$ and $g(\hat{\varepsilon}) = \phi(\hat{\varepsilon})/\sqrt{2\pi s^2(\hat{\varepsilon})}$, the stationary point $\hat{\varepsilon}^*$ satisfies $\hat{\varepsilon}^* = m^{-1}(L)$ and $\psi''(\hat{\varepsilon}^*) = [m'(\hat{\varepsilon}^*)/\hat{s}(\hat{\varepsilon}^*)]^2$, so that

$$f_{R_\pi}(L) \quad \sim \quad \frac{\phi(\hat{\varepsilon}^*)}{|m'(\hat{\varepsilon}^*)|}, \qquad N \to \infty. \tag{25}$$

The cumulative density is

$$F_{R_\pi}(L) \sim \int_0^L \frac{\phi[\hat{\varepsilon}^*(L')]}{-m'[\hat{\varepsilon}^*(L')]}\mathrm{d}L' = \Phi[-m^{-1}(L)]. \tag{26}$$

From eq. (22), VaR_q is just the inverse of the cumulative density function, *i.e.* for a given confidence level $0 \le q \le 1$, the Value-at-Risk, VaR_q, is found by setting $F_{R_\pi}(L) = q$ and solving for L:

$$L \quad = \quad m[-\Phi^{-1}(q)], \tag{27}$$

$$\Rightarrow \mathrm{VaR}_q \quad \sim \quad c\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) + \sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q)}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right], \tag{28}$$

as $N \to \infty$. □

Equation (28) is our main contribution for this paper. What are the errors associated with this asymptotic approximation? There are actually two contributions. One is associated with approximating the density of R_π, conditioned on $\hat{\varepsilon}$, with a Gaussian through the central limit theorem. The other is associated with applying Laplace's method to the integral in (24): see the higher order term in eq. (44). The first error appears as an extra term under the integral in (24): for large but finite N, $f_{R_\pi|\hat{\varepsilon}}$ would actually take the form of a Gaussian plus a small correction (Berry, 1941; Esseen, 1942). The second gives rise to an additive $O(N^{-1})$ term in all of eqs. (25)-(28). By comparing our results for VaR with Monte Carlo simulations in Fig. 2, we find that the dominant error term is in fact $O(N^{-1})$:

$$\mathrm{VaR}_q = c\sum_{i=1}^{N}\sum_{j=1}^{n} w_{ij}\Phi\left[\frac{\Phi^{-1}(PD_{ij}) + \sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q)}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right] + O(N^{-1}), \tag{29}$$

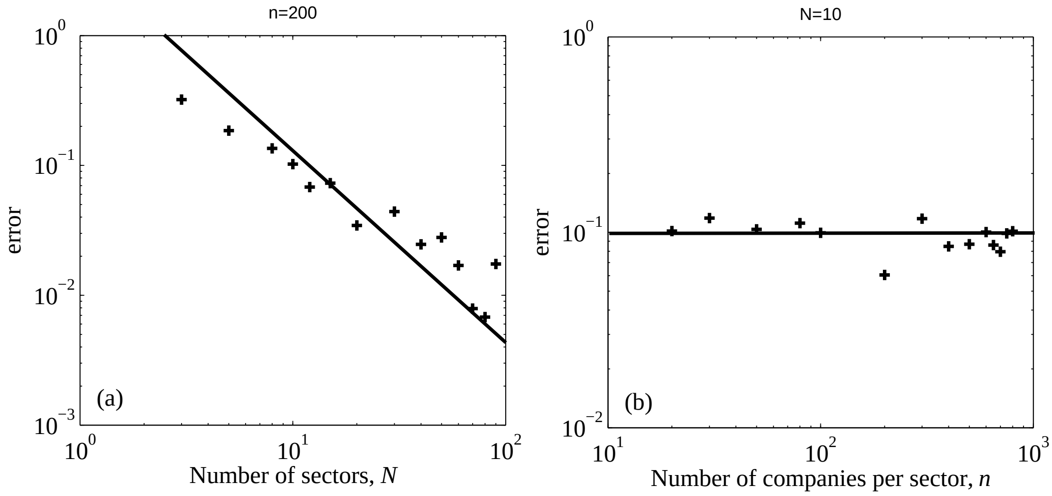

Figure 2: Error of analytic approximation (28), as defined by (30), scales as $O(N^{-1})$ when compared with Monte Carlo simulation and is independent of the number of companies per sector n. Solid line has slope -1 in (a) and 0 in (b). Parameters used were $c = 2$, $w_{ij} = 1/(Nn)$, $\beta_{ij} = 0.8$, $\theta_{ij} = -1.3$, $\rho_{ij} = 0.7$. Monte Carlo VaRs are found from 20,000 realizations.

as $N \to \infty$. Therefore, either the two contributions are of the same order or the error incurred by using Laplace's method is dominant. Eq. (29) is confirmed numerically by comparing the double sum with Monte Carlo simulations of the Value-at-Risk, $\mathrm{VaR}_q^{(num)}$. The error was measured using the infinity norm over $0 \le q \le 0.99$:

$$\text{error} = \max_{0\le q\le 0.99} \left| c \sum_{i=1}^{N} \sum_{j=1}^{n} w_{ij} \Phi \left[\frac{\Phi^{-1}(PD_{ij}) + \sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q)}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}} \right] - \mathrm{VaR}_q^{(num)} \right|. \tag{30}$$

The compared portfolios in Fig. 2 are completely homogeneous (and therefore somewhat artificial), but their purpose is to provide a benchmark result to deduce the scaling of the error. Note there is no scaling with n – the errors in Fig. 2(b) arise from N being finite.

Once the approximation to the VaR has been obtained, related quantities such as the expected shortfall and expected shortfall contribution are easily approximated as

$$\mathrm{ES}_q = \frac{1}{1-q}\int_q^1 \mathrm{VaR}_{q'}\mathrm{d}q', \tag{31}$$

$$\sim \sum_{i=1}^{N}\sum_{j=1}^{n}\frac{cw_{ij}}{1-q}\int_q^1 \Phi\left[\frac{\Phi^{-1}(PD_{ij})+\sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q')}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right]\mathrm{d}q', \tag{32}$$

$$\mathrm{ES}_{q,ij} = w_{ij}\frac{\partial \mathrm{ES}_q}{\partial w_{ij}}, \tag{33}$$

$$\sim \frac{cw_{ij}}{1-q}\int_q^1 \Phi\left[\frac{\Phi^{-1}(PD_{ij})+\sqrt{\rho_{ij}}\hat{\beta}_{ij}\Phi^{-1}(q')}{\sqrt{1-\rho_{ij}+\beta_{ij}^2\rho_{ij}}}\right]\mathrm{d}q', \tag{34}$$

as $N \to \infty$.

Now consider the case where the parameters β_{ij}, ρ_{ij} and θ_{ij} are only sector dependent: $\beta_{ij} = \beta_i$, $\rho_{ij} = \rho_i$ and $\theta_{ij} = \theta_i$ for the sector index i. Then all companies within a sector are statistically identical and one can consider a distribution of exposures that only depends on the sector. Defining $\hat{w}_i = \sum_{j=1}^n w_{ij}$, eq. (29) reduces to

$$\mathrm{VaR}_q = c\sum_{i=1}^{N}\hat{w}_i\Phi\left[\frac{\Phi^{-1}(PD_i)+\sqrt{\rho_i}\hat{\beta}_i\Phi^{-1}(q)}{\sqrt{1-\rho_i\hat{\beta}_i^2}}\right]+O(N^{-1}), \tag{35}$$

where $\sum_{i=1}^N \hat{w}_i = 1$ and $PD_i = \Phi(\theta_i)$; compare with eq. (6). It is easy to show that when $n = 1$ (recall that n does not have to be large for our approximations to be valid), eq. (35) is identical to the VaR for N firms coupled to a single factor, with rescaled systematic and idiosyncratic risks. In fact, eq. (1) implies that the asset return of each firm is given by

$$z_{i1} = \sqrt{r_i}\hat{\varepsilon} + \sqrt{1-r_i}\eta_i, \qquad r_i = \rho_{i1}\hat{\beta}_{i1}^2, \qquad i = 1,\ldots,N, \tag{36}$$

with $\sqrt{1-r_i}\eta_i \equiv \sqrt{\rho_{i1}}\beta_{i1}\varepsilon_1 + \sqrt{1-\rho_{i1}}\zeta_{i1}$ and $\eta_i \sim N(0,1)$. Application of Vasicek's formula to (36) then yields (35). The general case with n identical companies in each sector also collapses to the single-factor case because the first firm is representative of all firms in that sector, and its asset return is described by eq. (36). When the sectors are homogeneous, the hierarchical multi-factor model can in fact be treated as a single-factor model.

4 Results and Discussion

To validate our analytic approximation (28), we construct a proxy portfolio using exposure data taken from the call report of a large bank (JPM, 2014) with $N = 17$ "sectors" (see Table 1). In this example, the sectors correspond to different types of institution that borrow from the bank. For each of the sectors, we estimate sector default probabilities $\overline{PD}_i$ and variances of systematic risk $\bar{\rho}_i$. In this example, we also assume that the factor loadings are constant for every company. The $\overline{PD}_i$ are estimated by assuming that large government institutions and corporations are less likely to default than small companies and consumers. More accurate values could come from the credit rating of these entities (Crouhy et al., 2000).

From the sector parameters $\bar{\rho}_i$, $\overline{PD}_i$ and $\bar{w}_i$, we generate firm-level parameters by

$$\begin{aligned} \rho_{ij} &= \bar{\rho}_i + \delta\rho_{ij}, & (37)\\ \theta_{ij} &= \Phi^{-1}(\overline{PD}_i) + \delta\theta_{ij}, & (38)\\ \beta_{ij} &= \text{constant}, & (39)\\ w_{ij} &= \bar{w}_i/n, & (40) \end{aligned}$$

for $i = 1, \ldots, N$, $j = 1, \ldots, n$ with $\delta\rho_{ij} \sim N(0, 10^{-4})$ and $\delta\theta_{ij} \sim N(0, 10^{-4})$. Therefore, each of the 17 loans in Table 1 is subdivided equally into $n = 1000$ subloans, with corresponding exposure $\bar{w}_i/n$. At the sector level, this portfolio is quite typical in the sense that the distribution of exposures is "lumpy," with the portfolio being dominated by a few large loans, in this case to residential real estate (32%), US commercial and industrial companies (14%) and non-depository/other institutions (16%).

i	Loan Sector	Exposure (\$)	$\bar{w}_i$ (%)	$\overline{PD}_i$	$\bar{\rho}_i$
1	construction, land development	3,815	0.60	10^{-2}	0.23
2	farmland	211	0.03	10^{-2}	0.22
3	**1-4 residential properties**	203,246	**32.17**	10^{-2}	0.25
4	5+ residential properties	45,090	7.14	10^{-2}	0.21
5	nonfarm, non-residential	27,153	4.30	10^{-2}	0.27
6	commercial US banks	3,157	0.50	10^{-3}	0.15
7	banks in foreign countries	18,933	3.00	10^{-3}	0.13
8	agricultural loans	788	0.12	10^{-2}	0.12
9	**US commercial/industrial**	90,879	**14.39**	10^{-2}	0.18
10	non-US commericial/industrial	33,624	5.32	10^{-2}	0.28
11	credit cards	26,189	4.15	0.05	0.15
12	other revolving credit plans	2,584	0.41	0.05	0.17
13	automobile loans	41,517	6.57	0.05	0.21
14	other consumer loans	19,837	3.14	0.05	0.16
15	foreign governments	1,031	0.16	10^{-4}	0.12
16	US states and subdivisions	12,680	2.01	10^{-4}	0.17
17	**non-depository and other inst.**	101,000	**15.99**	10^{-2}	0.15
	Total	631,734	100.00		

Table 1: Proxy credit portfolio motivated by exposures taken from a JP Morgan call report. The three largest obligors are highlighted in bold face. Exposure dollar amounts are in millions.

Because the firm-level parameters are small perturbations of the sector-level parameters through eqs. (37)-(40), the correlation between two companies in sector i is approximately given by $\bar{\rho}_i$ and the correlation between companies in sector i and j ($i \neq j$) is approximately

$\sqrt{\bar{\rho}_i\bar{\rho}_j}\hat{\beta}_i\hat{\beta}_j$. The loading factors $\hat{\beta}_i$ essentially control cross-sector correlations. For the parameters in Table 1, when $\hat{\beta}_i = 0.95$, firms in different sectors are correlated at about 11 – 25%, when $\hat{\beta}_i = 0.87$, they are correlated at about 9 – 21% and when $\hat{\beta}_i = 0.6$, they are correlated at about 4 – 10%.

We now compare the Values-at-Risk predicted by eq. (28) with portfolio losses generated by drawing random variables defined by (1). Our Monte Carlo (MC) simulations use $n = 1000$ companies per sector and 1000 trials to simulate the value of the portfolio at some risk level q. In Figure 3, we see that the agreement is good providing the β_{ij} are not too large. As β_{ij} increases from 0.3 to 0.5 to 0.8, our analytic approximation (28) becomes worse, particularly for smaller values of q. The error bars for the Monte Carlo (MC) simulated VaR represent 99% confidence intervals. We see that for $\beta_{ij} \equiv 0.3$ and 0.5, the analytic solution is within the intervals for $q = 0.2, 0.4, 0.6$ and 0.8. For large factor loading $\beta_{ij} \equiv 0.8$, there is a significant departure from the MC simulations especially when $q \lesssim 0.6$.

The portfolio in Table 1 is fairly homogeneous in in terms of the systematic risk variance ρ_i. Many portfolios of interest are more inhomogeneous in that they contain a few companies or sectors whose defaults are very strongly correlated while the default of the majority of the companies is only weakly correlated. In Table 2, we sharply increase the value of $\bar{\rho}_i$ for large institutions in the portfolio. Again, we test the analytic VaR of eq. (28) against MC simulations when $\beta_{ij} = 0.3, 0.5$ and 0.8: see Figure 4 (note the β_{ij} in Table 2 are *not* used for these results; they are used in the next set of simulations described below). As in the first portfolio, the agreeement is good when $\beta_{ij} = 0.3$ or 0.5. In this portfolio, the firm-firm correlations have a wide range, spanning about 14% - 97% when $\beta_{ij} = 0.3$, 12% – 97% when $\beta_{ij} = 0.5$ and 6% – 97% when $\beta_{ij} = 0.8$. The strongest correlations are, of course, between companies in the same sector and obligors belonging to US states and subdivisions are the most strongly correlated in this portfolio. These institutions are responsible for the upper bound of ≈ 97% in the correlation matrix since $\rho_{16} = 0.97$ in Table 2.

In Fig. 5(a), we plot the VaR for the portfolio in Table 2, relax the constant β_{ij} assumption and instead generate them at the firm level through

$$\beta_{ij} = \bar{\beta}_i + \delta\beta_{ij}, \tag{41}$$

where $\bar{\beta}_i$ are taken from Table 2 and $\delta\beta_{ij} \sim \mathcal{N}(0, 10^{-4})$. We choose the loading factors $\bar{\beta}_i$ to be closer to 0 for loans with larger exposures. Hence the three largest loans are assigned $\bar{\beta}_i = 0.05$, loans with exposures between about 2% and 7% are assigned $\bar{\beta}_i = 0.1$ and the remaining loans are assigned values from 0.8 to 0.9. The rationale is that we wish to mimic a portfolio containing a few large loans whose defaults may be highly correlated with respect to global risk. Therefore we choose the loading factors so that the corresponding firms are more tightly coupled to $\hat{\varepsilon}$. Now, with β_{ij} stochastically generated through (41), the agreement with the MC simulated VaR is excellent for a large range of q values. The reason could be that even though there are 6 sectors (construction, farmland, commercial US banks, agriculture, revolving credit plans, foreign governments) that have large $\bar{\beta}_i$, the average factor loading across all companies in the entire portfolio is small and we have seen from Figs. 3 and 4 that our analytic solution performs better when firm factor loadings β_{ij} are small.

So far, our results have concentrated on a portfolio with $N = 17$ sectors. As discussed before, the value of N depends on how loans are classified. Is the approximation (28) still

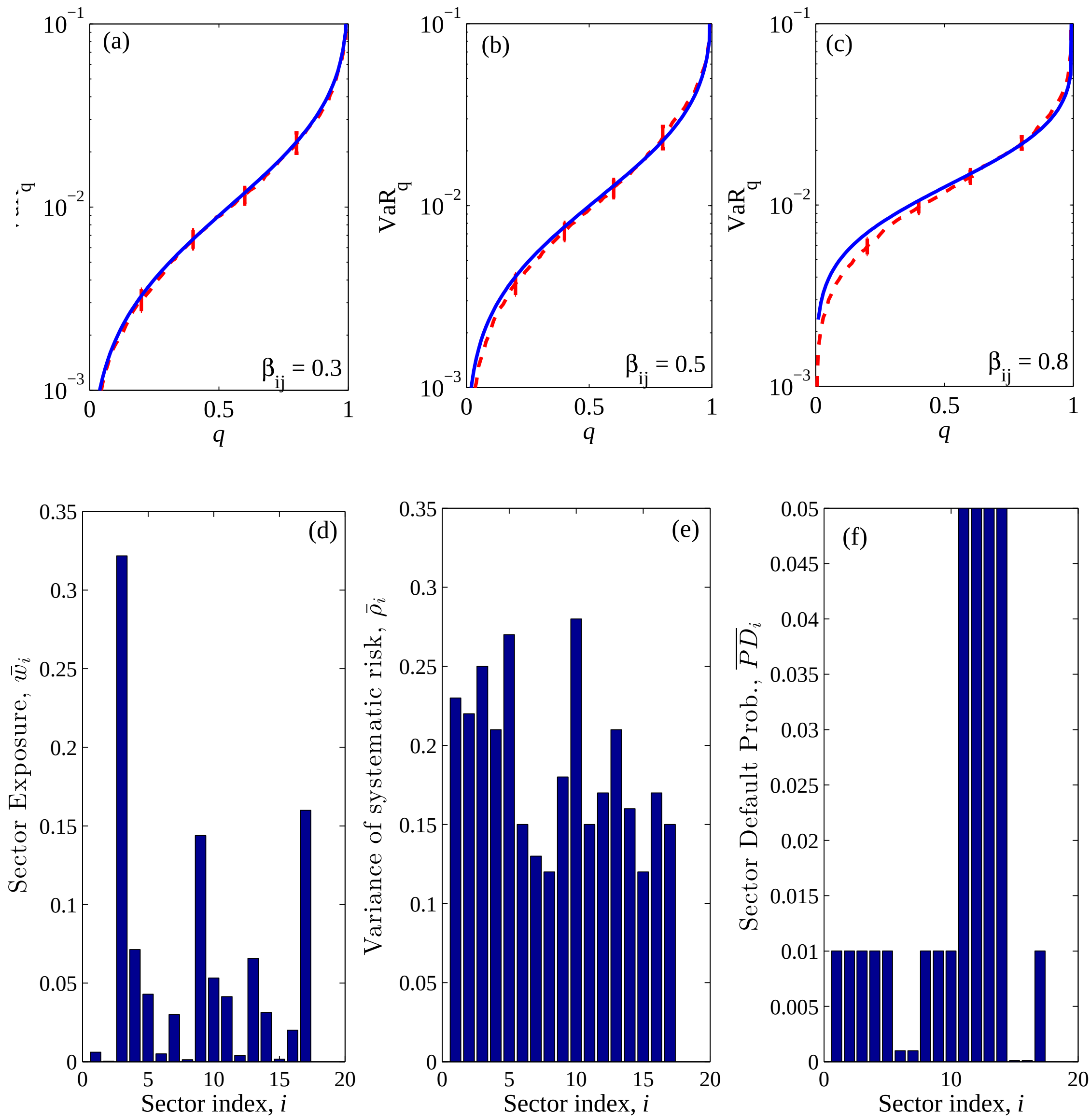

Figure 3: Comparison of Values-at-Risk (VaR) generated through Monte Carlo (MC) simulation (dashed red) and analytic approximation (solid blue) as determined by eq. (28). Error bars for the MC simulations are 99% confidence intervals derived through 50,000 Bootstrap samples. Parameters are taken from Table 1 with intra-sector correlations given by $\bar{\rho}_i$. Cross sector correlations are 11 – 25% for (a), 9 – 21% for (b) and 4 – 10% for (c). (d-f): Distribution of sector exposures, systematic risk variances and default probabilities.

i	Loan Sector	Exposure ($)	$\bar{w}_i$ (%)	$\overline{PD}_i$	$\bar{\rho}_i$	$\bar{\beta}_i$
1	construction, land development	3,815	0.60	10^{-2}	0.33	0.8
2	farmland	211	0.03	10^{-2}	0.32	0.9
3	**1-4 residential properties**	203,246	**32.17**	10^{-2}	0.35	0.05
4	5+ residential properties	45,090	7.14	10^{-2}	0.31	0.1
5	nonfarm, non-residential	27,153	4.30	10^{-2}	0.37	0.1
6	commercial US banks	3,157	0.50	10^{-3}	0.95	0.9
7	banks in foreign countries	18,933	3.00	10^{-3}	0.93	0.1
8	agricultural loans	788	0.12	10^{-2}	0.52	0.7
9	**US commercial/industrial**	90,879	**14.39**	10^{-2}	0.48	0.05
10	non-US commericial/industrial	33,624	5.32	10^{-2}	0.5	0.1
11	credit cards	26,189	4.15	0.05	0.15	0.1
12	other revolving credit plans	2,584	0.41	0.05	0.17	0.8
13	automobile loans	41,517	6.57	0.05	0.21	0.1
14	other consumer loans	19,837	3.14	0.05	0.16	0.1
15	foreign governments	1,031	0.16	10^{-4}	0.92	0.8
16	US states and subdivisions	12,680	2.01	10^{-4}	0.97	0.1
17	**non-depository and other inst.**	101,000	**15.99**	10^{-2}	0.85	0.05
	Total	631,734	100.00			

Table 2: A more strongly correlated proxy credit portfolio motivated by exposures taken from a JP Morgan call report. The three largest obligors are highlighted in bold face. Exposure dollar amounts are in millions.

i	Loan Sector	$\bar{w}_i$ (%)	$\overline{PD}_i$	$\bar{\rho}_i$	$\bar{\beta}_i$
1	Real Estate	44.25	10^{-2}	0.33	0.39
2	Depository Institutions and Banks	3.5	10^{-3}	0.94	0.5
3	Agricultural Loans	0.12	10^{-2}	0.52	0.7
4	Commercial/Industrial Loans	19.71	10^{-2}	0.49	0.075
5	Consumer Loans	14.27	0.05	0.32	0.45
6	Foreign Governments	0.16	10^{-4}	0.92	0.8
7	US States and Subdivisions	2.00	10^{-4}	0.97	0.1
8	Non-depository and other institutions	15.99	10^{-2}	0.85	0.05
	Total	100.00			

Table 3: Portfolio containing $N = 8$ sectors, based on a grouping together loans from Table 2.

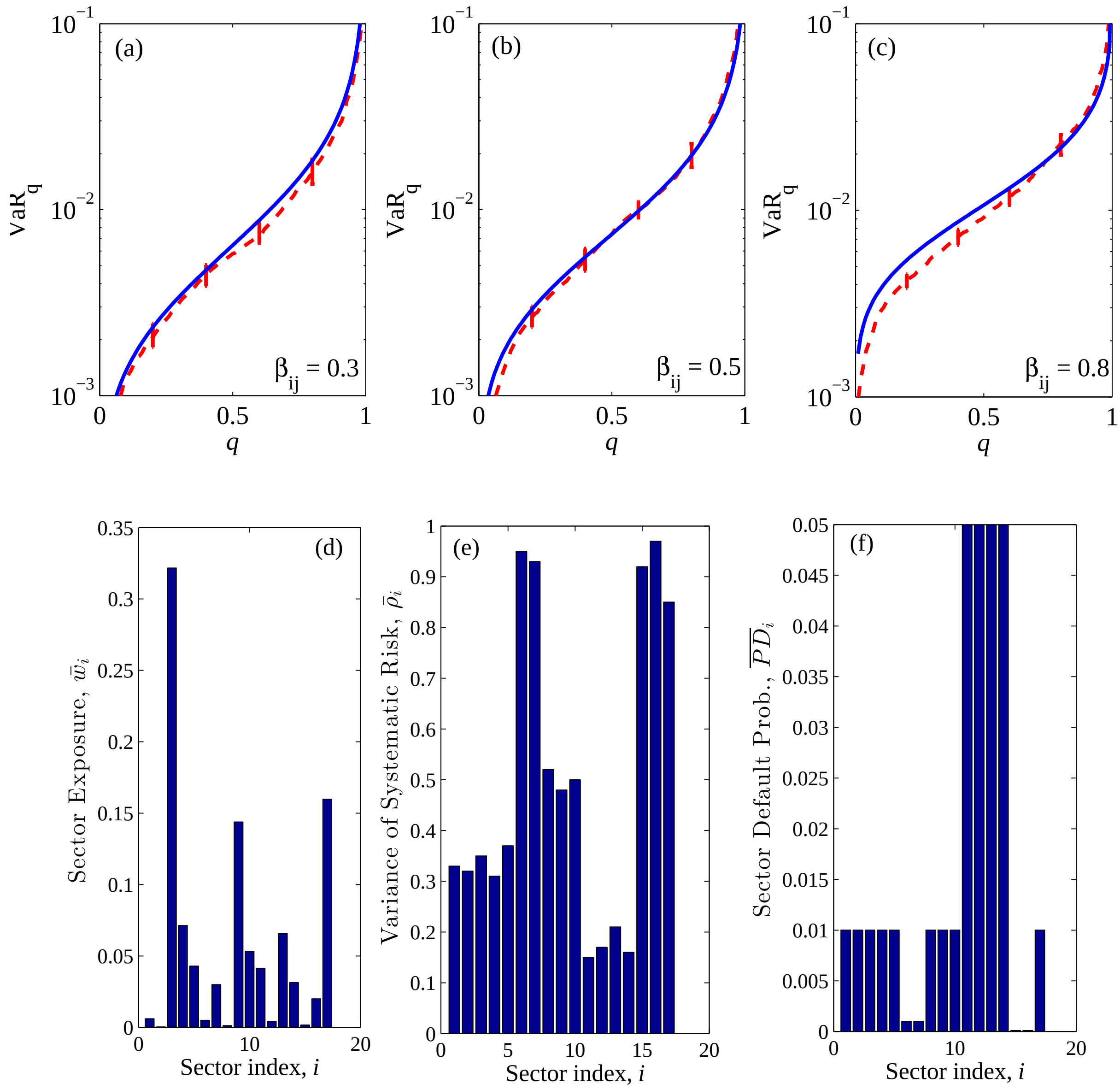

Figure 4: Comparison of Values-at-Risk (VaR) generated through Monte Carlo (MC) simulation (dashed red) and analytic approximation (solid blue) as determined by eq. (28). Error bars for the MC simulations are 99% confidence intervals derived through 50,000 Bootstrap samples. Parameters are taken from Table 2 with intra-sector correlations given by $\bar{\rho}_i$. Cross sector correlations are 14 – 87% for (a), 12 – 72% for (b) and 8 – 49% for (c). (d-f): Distribution of sector exposures, systematic risk variances and default probabilities.

accurate when N is reduced? In Table 3, we reduce N by grouping together loans that are economically similar. For example, we group commericial US bank loans and loans to foreign banks into a single "Depository Institutions and Banks" sector. This results in a new portfolio with $N = 8$ sectors, with some of the new sectors encompassing several of the old sectors in the $N = 17$ portfolio. The new sector exposures $\bar{w}_i$ are sums of the exposures in the old portfolio and the new $\overline{PD}_i$, $\bar{\rho}_i$ and $\bar{\beta}_i$ are averages of the parameters in the old portfolio. In Fig. 5(b), we see that although the analytic approximation (28) becomes worse for smaller values of q, it is still lies within the 99% confidence intervals for $q \gtrsim 0.4$. In particular, even though $N = 8$ is not "large," the agreement between the analytic solution and simulation results is still excellent for values of q close to 1.

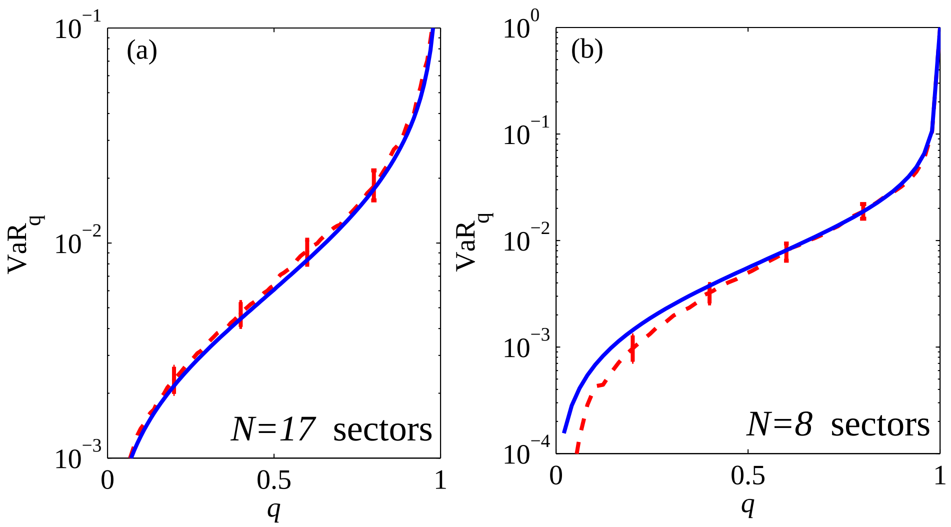

Figure 5: VaR for portfolios containing (a) $N = 17$ and (b) $N = 8$ sectors, corresponding to Tables 2 and 3. Solid blue curve indicates analytic solution (28), dashed red curve indicates Monte Carlo simulated values and error bars are 95% confidence intervals generated using 50,000 Bootstrap samples. Blue solid curve indicates analytic solution eq. (28).

We now compare the expected shortfall for MC simulated portfolios and the analytic approximation eq. (32). The analytic approximation is computed using a compound trapezoid rule with 2501, 5001 and 10001 abscissae for $q = 0.95, 0.90$ and 0.80 respectively. The MC expected shortfall ES_q is computed by finding the mean loss conditioned on the loss being larger than VaR_q:

$$\mathrm{ES}_q = E[R_\pi | R_\pi > \mathrm{VaR}_q]. \tag{42}$$

From Table 4, we see that our analytic approximation generally does a reasonable job in predicting the expected shortfall for portfolios with small β_{ij}. For portfolios 1 and 2 and 4, the relative error in ES is about 5-10%. When $\beta_{ij} = 0.8$, the ES disagree by up to 25%.

Finally, we try to formulate some guidelines for when our analytic approximation (28) is valid: see Table 5. This table provides a range of factor loadings β_{ij} and risk variances ρ_{ij} within which the agreement between the analytic solution (28) and Monte Carlo simulated

Portfolio 1: $\beta_{ij} = 0.3$	$q = 0.95$	$q = 0.90$	$q = 0.80$
ES_q (analytic)	0.1172	0.0795	0.0519
ES_q (MC)	0.1237	0.0848	0.0545
Portfolio 2: $\beta_{ij} = 0.5$			
ES_q (analytic)	0.1019	0.0720	0.0487
ES_q (MC)	0.1123	0.0800	0.0527
Portfolio 3: $\beta_{ij} = 0.8$			
ES_q (analytic)	0.0631	0.0497	0.0377
ES_q (MC)	0.0841	0.0611	0.0433
Portfolio 4: random β_{ij}			
ES_q (analytic)	0.1241	0.0825	0.0531
ES_q (MC)	0.1396	0.0934	0.0591

Table 4: Comparison of Expected Shortfall (ES) at level q (average loss in the worst $100(1-q)\%$ of cases). Portfolios 1,2,3 have $\beta_{ij} \equiv 0.3, 0.5, 0.8$ respectively and other firm parameters are generated through eqs. (37), (38) and (40). Portfolio 4 has β_{ij} generated through eq. (41). All values are correct to 4 decimal places.

VaR is good. Specifically, a checkmark in the Table indicates that the analytic solution lies within the 99% confidence intervals of the Monte Carlo solution at $q = 0.1, 0.3, 0.5$ and 0.7. The parameter values were generated using eqs. (37)-(40) with $\bar{\rho}_i = \bar{\rho}_0 + 0.05 Z_i$ where $Z_i \sim \mathcal{N}(0, 1)$. We set β_{ij} to be constant across all companies and sectors and the sector default probabilities $\overline{PD}_i = 0.0102$ for all i. The sector exposures $\bar{w}_i$ in eq. (40) are randomly generated and there are $n = 1000$ companies per sector. We find that for portfolios with 17 sectors or more, our analytic approximation agrees with the MC simulations providing the asset returns are not larger than about 0.3 and the loading factors β_{ij} are not larger than about 0.7.

5 Conclusions

In this paper, we computed the Values-at-Risk (VaR) and Expected Shortfalls (ES) for a bond portfolio under a hierarchical multi-factor model for the asset returns. Our main results are eqs. (28) and (32) which are analytic approximations to the portfolio's VaR and ES, given a risk level $0 \leq q \leq 1$. Our approximation to the VaR is written entirely in terms of easy-to-compute special and elementary functions and represent an economically intuitive extension of Vasicek's ASRF result to multiple sectors. It is much quicker to compute than predicting the VaR_q through many trials of a Monte Carlo (MC) simulation. Our formulas for the VaR_q give good approximations to the loss as predicted by Monte Carlo simulations when the sector factor loadings and systematic risk variances are not too close to 1. When the sector factor loadings are increased, we find that our approximations deviate from MC simulated VaR_q for small values of risk level q.

Our approximations are able to account for asset-return correlations among companies

β_{ij}	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\bar{\rho}_0 = 0.1$	✓		✓	✓		✓	
$\bar{\rho}_0 = 0.2$	✓	✓	✓	✓			
$\bar{\rho}_0 = 0.3$	✓		✓	✓			
$\bar{\rho}_0 = 0.4$				✓			
$\bar{\rho}_0 = 0.5$							

β_{ij}	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\bar{\rho}_0 = 0.1$	✓	✓	✓	✓	✓	✓	✓
$\bar{\rho}_0 = 0.2$	✓	✓	✓		✓	✓	✓
$\bar{\rho}_0 = 0.3$	✓	✓	✓	✓		✓	
$\bar{\rho}_0 = 0.4$					✓		
$\bar{\rho}_0 = 0.5$							

β_{ij}	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\bar{\rho}_0 = 0.1$	✓	✓	✓	✓	✓	✓	
$\bar{\rho}_0 = 0.2$	✓	✓	✓	✓	✓	✓	
$\bar{\rho}_0 = 0.3$	✓	✓	✓			✓	✓
$\bar{\rho}_0 = 0.4$							

Table 5: Comparison of analytic solution with Monte Carlo simulated VaR for $N = 13$ (top rows), $N = 17$ (middle rows) and $N = 23$ (bottom rows) sector portfolios. A check mark indicates that the analytic solution falls within the 99% confidence intervals for the simulated VaR at $q = 0.1, 0.3, 0.5, 0.7$.

at a global *and* sector level. We derived the formulas by using the central limit theorem and Laplace's method to approximate the loss distribution conditioned on the global risk when the number of sectors N is large, and then integrating over the global risk. The final analytic approximations have a similar mathematical structure to the ASRF but explicitly feature local and global factor loadings.

Although we have given some guidelines, in terms of model parameters, for when the analytic solution (28) may be accurate, quantifying and understanding its accuracy in terms of the correlation among firms is still an open question (the correlation matrix is easily found from eq. (1)). Certainly, our result is exact in the limit as $N \to \infty$, but for finite values of N, we do not know which correlation matrices give good agreement between (28) and simulation results, and which do not. We have found instances where the agreement is good and there is a wide range of correlation among firms, and instances where the agreement is poor and there is a narrow range of correlation among firms.

In summary, our work builds on current research in developing analytic and numerical tools to study credit portfolios. We hope that it will motivate further studies to build analytic approximations which could be accurate even for a very small number of sectors (say $N = 3$ or 4) and extreme parameter values.

6 Acknowledgements

The authors wish to thank William Morokoff and Saiyid Islam for many helpful comments on the manuscript and for introducing them to the credit portfolio problem during the Delaware 2012 Mathematical Problems in Industry workshop. They also thank Lou Rossi and David Edwards for organizing the workshop.

References

Consolidated Reports of Condition and Income for A Bank With Domestic and Foreign Offices – FFIEC 031, JP Morgan Chase Bank, National Association. Columbus, Ohio, December 2014. URL `www.fdic.gov/bank/statistical`.

M. J. Ablowitz and A. Fokas. *Complex Variables.* Cambridge University Press, 1997.

L. Andersen, J. Sidenius, and S. Basu. All your hedges in one basket. *RISK*, 16:67 – 72, 2003.

Basel Committee on Banking Supervision. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. *BCBS Publications*, 2006.

C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers – Asymptotic Methods and Perturbation Theory.* Springer, 2010.

A. C. Berry. The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49:122 – 136, 1941.

C. Bluhm, L. Overbeck, and C. Wagner. *Introduction to Credit Risk Modeling, 2nd Edition.* Chapman & Hall/CRC, 2010.

J. R. Bohn and R. M. Stein. *Active credit portfolio management in practice*, volume 384. Wiley, 2009.

X. Burtschell, J. Gregory, and J. Laurent. A comparative analysis of cdo pricing models. *The Journal of Derivatives*, 16(4):9–37, 2009.

P. Crosbie and J. Bohn. Modeling default risk. *KMV Corporation*, 1999.

M. Crouhy, D. Galai, and R. Mark. A comparative analysis of current credit risk models. *Journal of banking and finance*, 24:59–117, 2000.

M. Denault. Coherent allocation of risk capital. *Journal of Risk*, 4:7–21, 2001.

A. Erdelyi. *Asymptotic Expansions.* Dover, 1956.

C. G. Esseen. On the liapunoff limit of error in the theory of probability. *Arkiv för matematik, astronomi oc fysik*, A28:1 – 19, 1942.

M. B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, 12(3):199 – 232, 2003.

X. Huang, C. Oosterlee, and J. Van der Weide. Higher order saddlepoint approximations in the vasicek portfolio credit loss model. *Journal of Computational Finance*, 11:93–113, 2007.

J. Huh and A. Kolkiewicz. Computation of multivariate barrier crossing probability and its applications in credit risk models. *North American Actuarial Journal*, 12(3):263–291, 2008.

T. R. Hurd. Credit risk modeling using time-changed brownian motion. *International Journal of Theoretical and Applied Finance*, 12:1213, 2009.

J. Jensen. *Saddlepoint Approximations*. Oxford University Press, 1995.

H. Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, 1997.

P. H. Kupiec. A generalized single common factor model of portfolio credit risk. *The Journal of Derivatives*, 15(3):25–40, 2008.

D. X. Li. On default correlation: A copula function approach. *The Journal of fixed income*, 9:43, 2000.

D. J. Lucas. Default correlation and credit analysis. *Journal of Fixed Income*, 4:76 – 87, 1995.

R. Lugannani and S. Rice. Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in Applied Probability*, pages 475–490, 1980.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.

R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, 29:449 – 470, 1974.

F. W. J. Olver. *Asymptotics and Special Functions*. A. K. Peters/CRC Press, 1997.

M. Pykhtin. Multi-factor adjustment. *RISK*, 17, March 2004.

P. J. Schönbucher. Factor models: Portfolio credit risks when defaults are correlated. *The Journal of Risk Finance*, 3:45–56, 2001.

A. De Servigny and O. Renault. *Measuring and managing credit risk*. McGraw-Hill New York, 2004.

O. A. Vasicek. Probability of loss on loan portfolio. *KMV Internal Report*, 1987.

A Laplace's method for evaluating integrals

In our calculation, we will employ Laplace's method to approximate integrals. Here we briefly review this method. Laplace's method is a technique for asymptotically evaluating integrals of the form

$$I(k) = \int_a^b g(t)\exp[-k\psi(t)]\mathrm{d}t, \tag{43}$$

when $k \gg 1$. The method relies on the important fact that when k is large, most of the mass of the integrand will be located around a *stationary* point t^* where $\exp[-k\psi(t)]$ is maximal, or equivalently where $\psi(t)$ is minimal: $\psi'(t^*) = 0$, $\psi''(t^*) > 0$. When $t^* \in (a, b)$, we make the approximations $g(t) \approx g(t^*) + g'(t^*)(t - t^*) + \ldots$ and $\psi(t) \approx \psi(t^*) + \psi''(t^*)(t - t^*)^2/2 + \ldots$ to find that

$$I(k) \sim e^{-k\psi(t^*)}\sqrt{\frac{2\pi}{k\psi''(t^*)}}\left\{g(t^*) + O\left(\frac{1}{k}\right)\right\}, \qquad k \to \infty. \tag{44}$$

The first term on the right hand side of (44) is Laplace's approximation to $I(k)$. The second term in the series can be used to give an error estimate of the first term. A full account of Laplace's method and smoothness conditions required for g and ψ can be found in many texts such as Ablowitz and Fokas (1997); Olver (1997); Erdelyi (1956). An explicit form for the $O(k^{-1})$ term in (44) can be found in Bender and Orszag (2010).

B Matlab codes for Value-at-Risk evaluation

The Matlab code `GenerateMatrices.m` generates the model parameters w_{ij}, β_{ij}, θ_{ij} and ρ_{ij} for the hierarchical multi-factor portfolios in this paper. The code `MCsimulatedVaR.m` is used to simulate the Values-at-Risk using Monte Carlo simulation after calling `GenerateMatrices.m` at the command prompt. Finally, `AnalyticVaR.m` is used to compute the analytic solution as given by eq. (28). The codes can also be found on the primary author's website `http://udel.edu/~pakwing/MATLAB_codes/JCRcodes.txt`.

```
function [w,beta,theta,rho,wvec,betavec,thetavec,rhovec] = GenerateMatrices(n)
% generate "reasonable" model parameters w_{ij}, \beta_{ij}, theta_{ij}, \rho_{ij}
% for a large bank like JP Morgan for N=17 sectors and
% specified n (# companies per sector). The w_{ij} are based on call report
% data

% N     Loan type/sector
% 1     construction and land development
% 2     farmland
% 3     1-4 residential properties
% 4     5+ residential properties
% 5     nonfarm, non-residential
% 6     commercial US banks
```

```
% 7     banks in foreign countries
% 8     agricultural loans
% 9     US commercial/industrial loans
% 10 non-US commercial/industrial loans
% 11 Credit Cards
% 12 Other revolving credit plans
% 13 automobile loans
% 14 other consumer loans
% 15 foreign governments
% 16 US states and subdivisions
% 17 non-depository financial institutions & other

wvec = [
   0.006038934108343
   0.000334001336005
   0.321727182643328
   0.071374977443038
   0.042981697993143
   0.004997356482317
   0.029969892391418
   0.001247360439679
   0.143856433245638
   0.053224933278880
   0.041455739282673
   0.004090329157525
   0.065719115957033
   0.031400874418664
   0.001632016006737
   0.020071739054729
   0.159877416760852];

rhovec = [
    0.3300
    0.3200
    0.3500
    0.3100
    0.3700
    0.9500
    0.9300
    0.5200
    0.4800
    0.5000
    0.1500
    0.1700
```

```
    0.2100
    0.1600
    0.9200
    0.9700
    0.8500];

% rhovec = [
%     0.23
%     0.22
%     0.25
%     0.21
%     0.27
%     0.15
%     0.13
%     0.12
%     0.18
%     0.28
%     0.15
%     0.17
%     0.21
%     0.16
%     0.12
%     0.17
%     0.15
%     ];

thetavec = [
  -2.326347874040841
  -2.326347874040841
  -2.326347874040841
  -2.326347874040841
  -2.326347874040841
  -3.090232306167814
  -3.090232306167814
  -2.326347874040841
  -2.326347874040841
  -2.326347874040841
  -1.644853626951473
  -1.644853626951473
  -1.644853626951473
  -1.644853626951473
  -3.719016485455709
  -3.719016485455709
  -2.326347874040841];
```

```
betavec = [
    0.8000
    0.9000
    0.0500
    0.1000
    0.1000
    0.9000
    0.1000
    0.7000
    0.0500
    0.1000
    0.1000
    0.8000
    0.1000
    0.1000
    0.8000
    0.1000
    0.0500];

N=length(wvec);
w = zeros(N,n);
rho = zeros(N,n);
theta = zeros(N,n);
for i=1:N
    w(i,:) = wvec(i)/n;
    rho(i,:) = rhovec(i) + 0.01*randn(1,n);
    theta(i,:) = thetavec(i) + 0.01*randn(1,n);
    beta(i,:) = betavec(i) + 0.01*rand(1,n);
end

function [VaR,VaR_upper,VaR_lower,R_pi] = MCsimulatedVaR(rho,beta,theta,w,...
c,q,num_trials,confidence,qs)
%
% This function produces a Monte Carlo simulated VaR_q plot for the
% hierarchical multi-factor credit
% portfolio problem and generates error bars at values of q specified
% in the qs vector
%
% rho:     matrix of asset return variances
% beta:    matrix of factor loadings
% theta:   matrix of threshold default values s.t. prob default =
%          Phi(theta_{ij})
% w:       matrix of exposures
```

```
% c:       loss given default (LGD). It’s a scalar e.g. c=1.
% q:       vector of risk-values at which to evaluate VaR. Must be between
%          0 and 1 e.g. q = linspace(0,1,50)
%
% num_trials:             number of portfolios to generate e.g. num_trials = 5000
% confidence:             confidence values to create error bars. It’s a
%                         number between 0 (no confidence) and 1 (complete confidence)
% VaR:                    Values-at-Risk corresponding to q
% [VaR_upper,VaR_lower]: The upper and lower error bars for VaR,
%                         calculated using the bootstrap method
% qs:                     Values of q at which to calculate error bars e.g.
%                         qs = [0.2 0.4 0.6 0.8]
% R_pi:                   Portfolio loss for each of the num_trials trials

NNN = 50000; % number of bootstrap samples
beta_hat = sqrt(1-beta.^2);

% parameter values
[N,n] = size(rho); % N: number of sectors, n: number of companies per sector
R_pi = zeros(1,num_trials);
for i=1:num_trials
    R_pi(i) = one_draw(rho,beta,beta_hat,theta,w,c,N,n);
end

P = zeros(NNN,length(qs));
for i=1:NNN
    if mod(i,2000) == 0
        sprintf('Generating bootstrap samples: %d/%d',i,NNN)
    end
    % generate num_trials random integers from [1:num_trials]
    k = randi(num_trials,1,num_trials);
    bootstrap_sample = R_pi(k);
    P(i,:) = quantile(bootstrap_sample,qs);
end
alpha = (1-confidence)/2; % e.g. alpha = 0.05 for confidence_percentage = 0.9
VaR_lower = quantile(P,alpha);
VaR_upper = quantile(P,1-alpha);
VaR = quantile(R_pi,q); VaR(end) = 1;
semilogy(q,VaR,'r--','LineWidth',2);
hold on;

for i=1:length(qs)
  val(i) = (VaR_lower(i) + VaR_upper(i))/2;
  errorbar(qs(i),val(i),val(i)-VaR_lower(i),VaR_upper(i)-val(i),'r-','LineWidth',2);
```

```
end

axis([0 1 1e-3 1e-1]);
h=gca;
set(h,'FontSize',14,'FontName','Times');
end

function R_pi = one_draw(rho,beta,beta_hat,theta,w,c,N,n)
    % different rows of rho, beta etc. correspond to different sectors,
    % following the convention in the paper
    epsilon_hat = randn(1,1);
    epsilon = kron(randn(N,1),ones(1,n));
    Z1 = sqrt(rho) .* beta_hat*epsilon_hat;
    Z2 = sqrt(rho) .* beta .* epsilon;
    Z3 = sqrt(1-rho) .* randn(N,n);
    Z = Z1+Z2+Z3;
    LOSSES = c*(Z <= theta);
    R_pi = sum(sum( w.*LOSSES ));
end

function VaR = AnalyticVaR(rho,beta,theta,w,c,q)
% provides analytic solution for N sectors and n companies per sector
% rho = A(:,1); beta = A(:,2); theta = A(:,3); w = A(:,4);
beta_hat = sqrt(1-beta.^2);
[N,n] = size(rho);
for i=1:length(q)
    arg = ( theta + sqrt(rho).*beta_hat.*invPhi(q(i)) )./ (sqrt(1-rho+beta.^2.*rho));
    VaR(i) = c*sum(sum(w.* Phi(arg)));
end

semilogy(q,VaR,'b-','LineWidth',2); hold on;
xlabel('q','FontSize',14,'FontName','Times','FontAngle','italic');
ylabel('VaR_q','FontSize',14,'FontName','Times');
end


function out = Phi(x)
out = 1/2*(1+erf(x/sqrt(2)));
end

function out = invPhi(x)
out = sqrt(2)*erfinv(2*x-1);
end
```