GSMMC 2015, RENSSELAER POLYTECHNICAL INSTITUTE

Risky Business: Localized Factor Methods for Credit Risk Analysis

Nathan Gold, Madelyn Houser, Hsuan-Wei Lee, Dan Lu,
Matthew Moye, Xiaoqi Zhang, Yingxiang Zhou
Supervised by Dr. Pak-Wing Fok

June 19, 2015

Abstract

Credit risk evaluation remains an important process for financial companies to remain profitable. Many methods are utilized to determine portfolio risk. This paper analytically verifies the Vasicek Asymptotic Single Risk Factor (ASRF) model of portfolio risk and compares the result to numerical simulations. The authors then explore extensions of the single factor model to a model that includes global, sector, and company-specific factors.

1 INTRODUCTION OF THE PROBLEM

A principal concern of the financial services industry is that of properly managing risk. Financial institutions including banks, insurance companies, and fund managers maintain credit portfolios - collections of bonds, mortgage backed securities, and various other loans - that have an inherent amount of default risk associated with them. The need to properly quantify the risk arising from credit portfolios has given rise to a range of models that aim to give risk managers a realistic expectation of their losses.

Credit portfolio returns on investment typically follow a heavily-skewed distribution. The up-side returns are limited (with high probability) while the down-side is potentially very large. The main risk of credit portfolios is that a debtor will default on their loan obligation, resulting in the lender possibly losing the entire amount of capital invested. Default is generally assumed to be a rare event that occurs with low probability. There have been, however, times during which the probability of default has been much higher than normally anticipated, such as the US Subprime Mortgage Crisis of 2007-2009. Thus, it is important to understand the risk associated with credit portfolios.

When modeling the distribution of returns for assets in a credit portfolio, the organization of companies in the global economy should be taken into consideration. Companies are typically organized by sector, such as technology, manufacturing, oil and gas, etc. In each sector, company performance is correlated to a large extent with company's share prices following similar trends to one another (Figures 1.1 and 1.2). Company performance within the economy is also correlated, but less so than within-sector correlation (Figure 1.3) . Thus, company defaults are not independent events because of correlations between asset returns. A representative model of the distribution of returns must incorporate the correlations of debtors within sectors as well as in the global economy.

We focus on a particular class of credit risk models known as a factor models. We first examine single factor model introduced by Vasicek [1]. We confirm Vasicek result using the Asymptotic Single Risk Factor and Laplace's method to determine the probability density function of the asset returns. We also compute the loss distribution and the VaR_q, an essential metric for practioners. We also extend the model to a localize multi-factor factor model for the global economy, sectors, and company-specific factors. Our model is compared to numerical simulations.

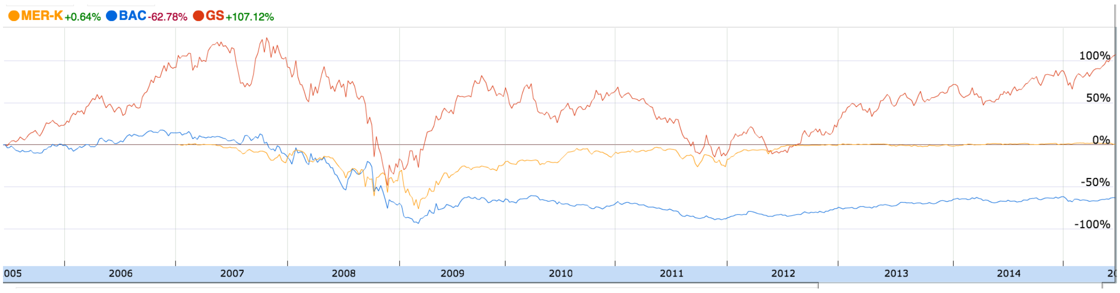

Figure 1.2: Bank of America, Merrill-Lynch, and Goldman-Sachs Historical Share Prices

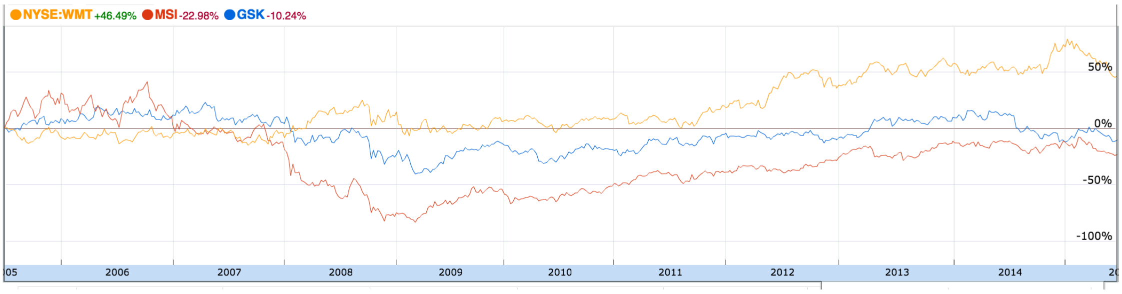

Figure 1.3: Wal-mart, Motorola, and GSK Historical Share Prices

Figure 1.1: Ford, General Motors, and Toyota Historical Share Prices

2 Analytical Method

We start with the single factor Vasicek model,

$$z_i = a\hat{e} + b\epsilon_i$$

where z_i is the asset level for company i, $\hat{e}$ is a global economic factor, and ϵ_i is the company specific idiosyncratic default risk. Additionally, $a = \sqrt{\rho}$, $b = \sqrt{1-\rho}$ where ρ is the correlation of assets between any two companies. Note that $a^2 + b^2 = 1$.

To calculate the loss distribution of the portfolio, we must calculate the cumulative distribution function (CDF) of the return of the portfolio. Define the return as

$$R(T) = \sum_{i=1}^{N} w_i R_i(z_i, T)$$

where T is the time horizon, w_i is the weight of the ith loan as a proportion of the entire portfolio (taken as $1/N$), and $R_i(z_i, T)$ is the loss on investment for the ith company in the portfolio.

We start by calculating the return for each company in the portfolio. Because $\hat{\epsilon} \sim N(0,1)$ and $\epsilon_i \sim N(0,1)$, then the distribution of assets given fixed economic state (i.e. for fixed $\hat{\epsilon}$)

$$z_i|\hat{\epsilon} \sim N(a\hat{\epsilon}, b^2)$$

Then the loss distribution for each i company in the portfolio

$$R_i(z_i|\hat{\epsilon}, T) = \begin{cases} 0 & : z_i > \theta|\hat{\epsilon} \\ c & : z_i \leq \theta|\hat{\epsilon} \end{cases}$$

where c is the amount an investor loses given company i's default and θ_i is the threshold of loss for the ith company in the portfolio (i.e. should the assets for company i fall below θ_i, we would expect a default on the loan). Recall that we want to find the total loss distribution for the portfolio containing N companies, hence for a given time horizon, T, we want to find

$$R(T) = \sum_{i=1}^{N} \frac{1}{N} R_i(z_i, T)$$

Now let the probability of default conditioned on $\hat{\epsilon}$ be defined as

$$p(\hat{\epsilon}) = P(z_i < \theta|\hat{\epsilon}) = \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi} b} \exp\left\{-\frac{(t - a\hat{\epsilon})^2}{2b^2}\right\} dt$$

Letting the portfolio loss be defined as the number of defaults over the number of loans in the portfolio, k/N, and using the law of total probability we have

$$P\left(R = \frac{k}{N}\right) = P(\text{k defaults}) = \int_{-\infty}^{\infty} \binom{N}{k} p(\hat{\epsilon})^k (1 - p(\hat{\epsilon}))^{N-k} g(\hat{\epsilon}) d\hat{\epsilon}$$

Let $y = \frac{k}{N}$. Then, approximating the binomial distribution with a normal distribution, we have the probability of default falling within a certain range given by

$$P(y \leq Y \leq y + \delta y) = \int_{-\infty}^{\infty} \frac{\sqrt{N}}{\sqrt{2\pi p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}} \exp\left\{\frac{-N(y - p(\hat{\epsilon}))^2}{2p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}\right\} g(\hat{\epsilon}) d\hat{\epsilon}$$

where $g(\hat{\epsilon})$ is the density function.

Letting

$$\phi(\hat{\epsilon}) = \frac{(y - p(\hat{\epsilon}))^2}{2p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}$$

then we have

$$P(\text{k defaults}) = \int_{-\infty}^{\infty} \frac{\sqrt{N}}{\sqrt{2\pi p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}} \exp\left\{-N\phi(\hat{\epsilon})\right\} g(\hat{\epsilon})\, d\hat{\epsilon}$$

We can use Laplace's method for evaluating integrals. Laplace's method states that if some function ψ has a minimum at t^* then

$$\int_a^b g(t) \exp\left\{-k\psi(t)\right\} dt \sim g(t^*) \exp\left\{-k\psi(t^*)\right\} \sqrt{\frac{2\pi}{\psi''(t^*)k}}$$

where $\psi'' \geq 0$ and $\psi'(\epsilon^*) = 0$.

Now let

$$\alpha(\hat{\epsilon}) = \frac{\sqrt{N}}{\sqrt{2\pi p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}} g(\hat{\epsilon}) = \frac{\sqrt{N}}{\sqrt{2\pi p(\hat{\epsilon})(1 - p(\hat{\epsilon}))}} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-\hat{\epsilon}^2}{2}\right\}$$

So we have

$$P(\text{k defaults}) = \int_{-\infty}^{\infty} \alpha(\hat{\epsilon}) \exp\left\{-N\phi(\hat{\epsilon})\right\} d\hat{\epsilon}$$

and we can use Laplace's method.

We first make a simplification for $p(\hat{\epsilon})$ by letting $m = \frac{t - a\hat{\epsilon}}{b}$.Then

$$p(\hat{\epsilon}) = \frac{1}{b} \int_{-\infty}^{\frac{\theta - a\hat{\epsilon}}{b}} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-m^2}{2}\right\} b\, dm = \Phi\left(\frac{\theta - a\hat{\epsilon}}{b}\right),$$

where Φ denotes the cumulative distribution function for standard normal distribution $N(0,1)$.

First note that for all $\hat{\epsilon}$, $\phi(\hat{\epsilon}) \geq 0$ so the minimum of $\phi(\hat{\epsilon})$ occurs when $y = p(\hat{\epsilon})$. Let $\hat{\epsilon} = \epsilon^*$ be the location of the minimum of $\phi(\hat{\epsilon})$. Then,

$$\epsilon^* = p^{-1}(y) = \frac{\theta - b\Phi^{-1}(y)}{a}$$

Since we have a minimum, ϵ^*, by the Laplace approximation, we have

$$I(N) = \alpha(\epsilon^*) \exp\left\{-N\phi(\epsilon^*)\right\} \sqrt{\frac{2\pi}{N\phi''(\epsilon^*)}}$$

We then compute $\phi''(\hat{\epsilon})$ using a computer algebra system and evaluate at ϵ^* for which the only non-zero term becomes

$$\phi''(\epsilon^*) = \frac{p'(\epsilon^*)^2}{p(\epsilon^*)(1 - p(\epsilon^*))}$$

Note that

$$p'(\hat{\epsilon}) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{\theta - a\hat{\epsilon}}{b}\right)^2/2} \left(-\frac{a}{b}\right)$$

As $\Phi^{-1}(y) = \frac{\theta - a\epsilon^*}{b}$ we have

$$p'(\epsilon^*) = -\frac{1}{\sqrt{2\pi}} \frac{a}{b} e^{-[\Phi^{-1}(y)]^2/2}$$

Then evaluating the Laplace approximation, we have

$$I(N) \sim \frac{1}{\sqrt{2\pi y(1-y)}} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{\theta - b\Phi^{-1}(y)}{a}\right)^2/2} \cdot \frac{\sqrt{2\pi y(1-y)}}{\sqrt{N}\frac{1}{\sqrt{2\pi}}\frac{a}{b} e^{-[\Phi^{-1}(y)]^2/2}}$$

So

$$I(N) \sim \frac{b}{a\sqrt{N}} \exp\left\{ \frac{[\Phi^{-1}(y)]^2 - \left[\frac{\theta - b\Phi^{-1}(y)}{a}\right]^2}{2} \right\}$$

Therefore, substituting back to the original correlation ρ, the density for the loss L is given by

$$f_L(y) = \sqrt{\frac{1-\rho}{\rho}} \exp\left\{ \frac{1}{2}[\Phi^{-1}(y)]^2 - \frac{[\theta - \sqrt{1-\rho}\Phi^{-1}(y)]^2}{2\rho} \right\}$$

We thus recover the result of Vasicek asymptotic single factor model [1].

Now we compute the CDF

$$f_L(y) = \sqrt{\frac{1-\rho}{\rho}} \exp\left\{ \frac{1}{2}\left(\Phi^{-1}(y)\right)^2 - \frac{1}{2\rho}\left]\theta - \sqrt{1-\rho}\Phi^{-1}(y)\right]^2 \right\}$$

Letting

$$u = \sqrt{1-\rho}\Phi^{-1}(y) - \theta \Rightarrow y = \Phi\left(\frac{u+\theta}{\sqrt{1-\rho}}\right)$$

$$y : 0 \to x \Rightarrow u : -\infty \to \sqrt{1-\rho}\Phi^{-1}(x) - \theta$$

Then

$$\begin{aligned} F_L(x) = P\{L < x\} &= \int_0^x \sqrt{\frac{1-\rho}{\rho}} \exp\left\{ \frac{1}{2}\left(\frac{u+\theta}{\sqrt{1-\rho}}\right)^2 - \frac{1}{2\rho}u^2 \right\} dy \\ &= \int_{-\infty}^{\sqrt{1-\rho}\Phi^{-1}(x)-\theta} \sqrt{\frac{1-\rho}{\rho}} \exp\left\{ \left(\frac{u+\theta}{\sqrt{1-\rho}}\right)^2 /2 - \frac{u^2}{2\rho} \right\} \frac{1}{\sqrt{2\pi}} \exp\left\{ \left(\frac{u+\theta}{\sqrt{1-\rho}}\right)^2 /2 \right\} \frac{dy}{\sqrt{1-\rho}} \end{aligned}$$

Letting $v = \frac{u}{\sqrt{\rho}}$ and simplifying terms, we have

$$F_L(x) = \int_{-\infty}^{\frac{\sqrt{1-\rho}\Phi^{-1}(x)-\theta}{\sqrt{\rho}}} \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv$$
$$= \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1}(x)-\theta}{\sqrt{\rho}}\right)$$

To find VaR_q, we compute the

$$\theta + \sqrt{\rho}\Phi^{-1}(q) = \sqrt{1-\rho}\Phi^{-1}(x)$$

$$\Phi^{-1}(x) = \frac{\theta + \sqrt{\rho}\Phi^{-1}(q)}{\sqrt{1-\rho}}$$

$$F^{-1}(x) = VaR_q = \Phi\left(\frac{\theta + \sqrt{\rho}\phi^{-1}(q)}{\sqrt{1-\rho}}\right)$$

3 Numerical Results

In Matlab, we make use of pseudo-random number generators and the statistics toolbox to numerically explore the nature of the portfolio loss. The parameters of the model, i.e. sector factor loadings, global factor loadings, and variance of system risk, are passed towards a function to generate draws of the i.d. random variables in a credit risk model. Our Monte Carlo method performs these trial by running many times (e.g. 100,000) with the total loss generated for each run. The sectors and companies, as well as any known default rates, can be modified by a user. We use the code to generate the PDF,CDF, VaR_q, and Expected Shortfall of a credit portfolio. Visualizing the VaR_q is quite easy in Matlab, and, given the monotonic nature of the CDF, the VaR_q is merely a reversal of the CDF.

To compute the Expected Shortfall as a function of q, rather than computing the full integral, one can use the probabilistic interpretation to develop an easy algorithm.

$$\text{ES} = E[L|L \geq VaR_q]$$

For each draw from our set, take the VaR_q. Consider only the set of losses greater than the VaR_q, and compute the mean.

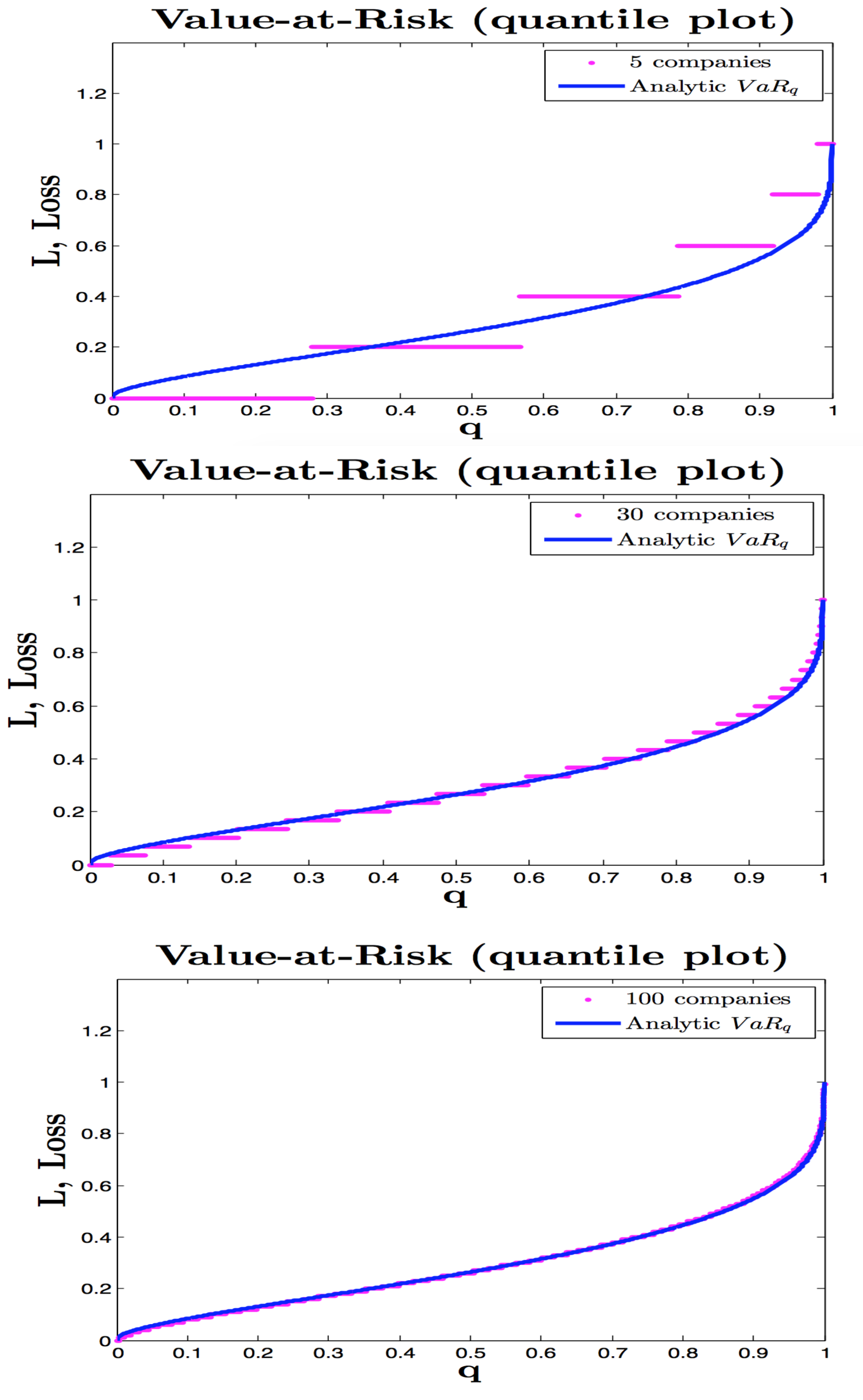

Figure 3.1: We consider the Vasicek model (single-factor) given a fixed $\rho = .2546$, $\theta = -.5447$, $c = 1$, and $w = \frac{1}{N}$. We enumerate the simulations over increasing number of companies, N. The results compared to the analytical result ($N \to \infty$) are given.

Note that the single factor model requires the assumption of a large number of companies in the portfolio. The above graphs demonstrate that the simulation results converge to the analytic VaR_q as the number of companies increases. Note that the stair-case nature of the numerical solution results from the discretization of the number of companies. As the number of companies increase to infinity, we have a probability of default ranging (continuously) between 0 and 1.

4 Multi-factor extension

We further explored the extension of the single factor model to a more complex economy in which the jth company's assets are correlated to the global economy ($\hat{\epsilon}$), the company's sector (ϵ_i), and the idiosyncratic factor for the jth company (ζ_{ij}).

$$z_{ij} = \sqrt{\rho_i}\left(\hat{\beta}_i\hat{\epsilon} + \beta_i\epsilon_i\right) + \sqrt{1-\rho_i}\zeta_{ij}$$

We use the same process as the single factor model to derive an analytic solution to the multi-risk model.

First for N companies in each of the M sections, we apply classic central limit theorem in each sector i, and get the density of loss in sector i conditioned on ϵ_i and $\hat{\epsilon}$

$$D_{R_i|\epsilon_i,\hat{\epsilon}} = \frac{1}{\sqrt{\frac{2\pi p(\epsilon_i,\hat{\epsilon}(1-p(\epsilon_i,\hat{\epsilon})}{M}}}\exp\left\{-\frac{[x-p(\epsilon_i,\hat{\epsilon})/M]^2}{\frac{2p(\epsilon_i,\hat{\epsilon})(1-p(\epsilon_i,\hat{\epsilon}))}{M}}\right\}.$$

Use the law of total probability, integrating over ϵ_i,

$$D_{R_i|\epsilon_i} = \int_{-\infty}^{\infty} D_{R_i|\epsilon_i,\hat{\epsilon}}(x)\frac{e^{-\epsilon_i^2/2}}{\sqrt{2\pi}}\,d\epsilon_i.$$

On the basis of

$$\mu(\hat{\epsilon}) = \frac{1}{M}\sum_{i=1}^{\infty}\mu_i(\hat{\epsilon}),$$

and

$$\sigma^2(\hat{\epsilon}) = \frac{1}{M^2}\sum_{i=1}^{\infty}\sigma_i^2(\hat{\epsilon}),$$

we use Lindeberg central limit theorem on all the sectors, which are independent but not identically distributed, and it yields the density of global loss conditioned on $\hat{\epsilon}$

$$D_{R|\hat{\epsilon}}(x) = \frac{1}{\sqrt{2\pi\sigma^2(\hat{\epsilon})}}\exp\left\{-\frac{[x-\mu(\hat{\epsilon})]^2}{2\sigma_i^2(\hat{\epsilon})}\right\}.$$

Integrating over the global factor $\hat{\epsilon}$, the density of the global (unconditioned) loss is

$$D_R(x) = \int_{-\infty}^{\infty} D_{R|\hat{\epsilon}}(x)\frac{e^{-\hat{\epsilon}^2/2}}{\sqrt{2\pi}}\,d\hat{\epsilon}.$$

If we define

$$s(\hat{\epsilon}) = \sigma^2(\hat{\epsilon})/M$$

the asymptotic solution is given by using Laplace method again:

$$D_R(x) \sim \sqrt{M} g(\epsilon^*) e^{-M\phi(\epsilon^*)} \frac{\sqrt{2\pi}}{M\phi''(\epsilon^*)} \frac{1}{2\pi}$$

$$D_R(x) \sim \sqrt{\frac{M}{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-\hat{\epsilon}^2/2}}{s(\hat{\epsilon})} e^{-M(x-\mu(\hat{\epsilon}))^2/2s^2(\hat{\epsilon})} \, d\hat{\epsilon}$$

$$= e^{-\epsilon^*(x)/2} \Bigg/ \left[\frac{1}{M} \sum_{i=1}^{M} \exp\left\{ -\frac{1}{2} \left(\frac{\theta_i - \sqrt{\rho_i}\hat{\beta}_i \epsilon^*}{\sqrt{1-\rho_i+\rho_i\beta_i^2}} \right)^2 \right\} \cdot \frac{\sqrt{\rho_i}\hat{\beta}_i}{\sqrt{1-\rho_i+\rho_i\beta_i^2}} \right]$$

where

$$\epsilon^*(x) = \mu^{-1}(y)$$

and

$$\mu(y) = \frac{1}{M} \sum_{i=1}^{M} \Phi\left(\frac{\theta_i - \sqrt{\rho_i}\hat{\beta}_i y}{\sqrt{1-\rho_i-\rho_i\beta_i^2}} \right).$$

We also examine the problem numerically. We numerically calculate the probability density function and expected shortfall of the loss on a portfolio given global, sector, and company-specific idiosyncratic risk.

The figures below show the numerical solutions to the full localized multi-factor model.

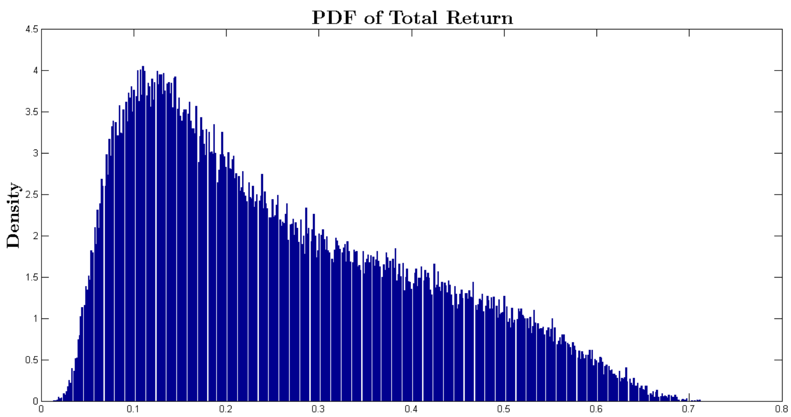

Figure 4.1: PDF of the full localized multi-factor risk model, c = .8, ρ is the correlation matrix uniformly drawn

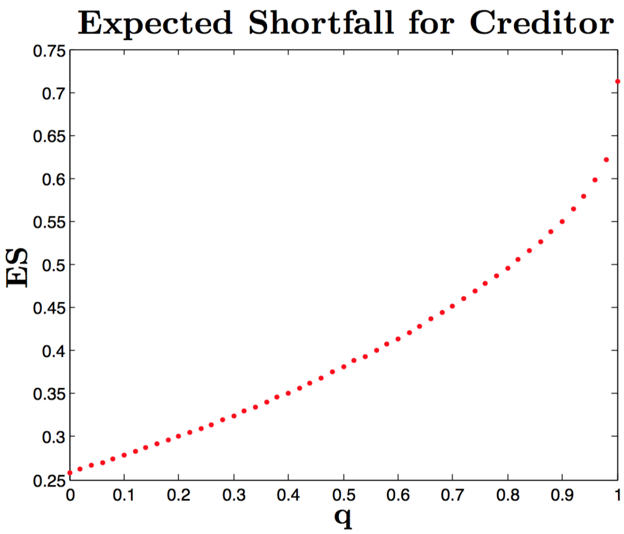

Figure 4.2: Expected Shortfall of the full localized multi-factor risk model

5 Conclusions and Future Work

We derive and verify, both analytically and numerically, the Vasicek single-factor risk model using asymptotic methods. Additionally, we explore the localized multi-factor model and generate partial analytic results. We perform numerical simulations on the multi-factor model to determine the loss distribution and expected shortfall of the full model. In the future, the authors would like to develop a closed-form solution to the localized multi-factor model.

References

[1] O.A. Vasicek, *Probability of Loss on Loan Portfolio*, KMV Internal Report (1987).

[2] X. Huang, C. Oosterlee and J. Van der Weide, *Higher order saddle-point approximations in the Vasicek portfolio credit loss model*, J. Computational Finance, 11: 93-113 (2007).

[3] M. Pykhtin, *Multi-factor Adjustment*, RISK 17 (2004).