Basel II Credit Loss Distributions under Non-Normality

Enrique Batiz-Zuk,[1] George Christodoulakis and Ser-Huang Poon[2]

Abstract

In the context of Vasicek (1987, 2002) single factor model, we examine the impact of skewness and excess kurtosis in the asset return process on the shape of the credit loss distribution and, consequently, over the Basel II requirements. We use Skew Normal and Skew Student's *t* densities to develop a Maximum Likelihood estimator of the credit loss density for aggregate charge-off rates published by the Federal Reserve Board for ten U.S. sectors. We show that, the non-gaussian modelling of the common factor provides a better characterization than its Gaussian counterpart, and has a significant impact on the capital requirement depending on the sign and magnitude of the skew-related coefficient. On the other hand, the non-gaussian modelling of the idiosyncratic factor does not provide a significantly better characterization than the Gaussian base case. The latter could be due to the fact that the sector portfolios are large and the idiosyncratic component has been fully diversified away.

JEL Classification: C13,C16, C22, G21, G32

Keywords: Vasicek loan loss distribution, single factor model, Basel II, non-Gaussian distributions , Skew Normal distribution, Skew Student's *t* distribution.

Acknowledgements:
The authors wish to thank Professor Adelchi Azzalini for his helpful comments and guidance on estimating the Skew Normal and the Skew Student's *t* distributions.

---

[1] A set of Matlab functions to compute the Skew Student's *t* distribution, density, quantiles, cumulants and a pseudo *ST* random number generator were ported from Azzalini's Skew Student's *t* R library. The Matlab functions of these routines, created by the first author, are now available on Azzalini's web page (http://azzalini.stat.unipd.it/SN/index.html#lib-sn).

[2] All authors are at Manchester Business School, University of Manchester, Crawford House, Oxford Road, Manchester M13 9PL, United Kingdom, Tel: +44 161 306 6404 (secretary), Fax: +44 161 275 4023, Email: Enrique.Batiz(at)postgrad.mbs.ac.uk, George.Christodoulakis(at)mbs.ac.uk, Ser-Huang.Poon(at)mbs.ac.uk.

Basel II Credit Loss Distribution under Non-Normality

1. Introduction

It is a regulatory requirement that financial institutions should reserve sufficient capital because of their exposure to credit and other risks. Whilst the adequacy of such reserves is crucial for their survival as well as the systemic financial stability, the ongoing credit crisis has placed a serious doubt on the way these reserves are calculated. Pillar I of Basel II (2004) provides the regulatory framework for determining bank capital requirements for taking credit risk. Under this regulatory framework, a bank may choose the internal ratings-based approach that utilizes risk weights derived implicitly from Vasicek's (1987, 2002) single factor model (SFM). In the Vasicek's model, changes to asset value are driven by a common and an idiosyncratic risk factors both of which are assumed to be Gaussian. However, the imposed Gaussianity assumption can be a poor proxy of the true and unobservable distribution, leading to a higher capital charge when the right tail of the distribution is underrepresented and a lower capital charge when the left tail is underrepresented and vice cersa when the right tail is overrepresented. The latter is particularly important as risk is not adequately covered. In this paper we relax the Gaussianity assumption and estimate with maximum likelihood generalized Vasicek credit loss distributions that are based on asset processes that feature skewness and excess kurtosis. Our data concern quarterly charge-off rates in ten US sectors from 1985 to 2007. Our findings provide overwhelming evidence in favour of non-normality and leads to significantly different capital charge calculation as compared to those in Basel II.

Since the publication of Vasicek (1987), there have been a number of theoretical extensions (see Batiz, Christodoulakis and Poon (2008) for a comprehensive survey) for the credit loss distribution. In practice, the common factor is unobservable and there is no empirical methodology available to study and test the Gaussian assumption. This departure from Normality could exert a large impact on the loan portfolio loss distribution and, thus, the regulatory capital charges (see Schönbucher (2001)). In this paper we model and assess the impact on Vasicek's capital charges due to non-Gaussian common or idiosyncratic factors separately. We consider Skew Normal (*SN*) and Skew Student's *t* (*ST*) as alternatives to Gaussian. These two non-Gaussian densities encompass the normal as a special case and both have the property of being analytically tractable. Moreover, both are very flexible for controlling the amount of skewness and excess kurtosis in the distribution. The *SN* achieves this, to a moderate degree, through a single additional parameter. The *ST*, which includes *SN* as a limiting case, provides a much greater flexibility over the degree of skewness and excess kurtosis through two additional parameters (see Azzalini (2005)), over the Normal.

To compare Vasicek's Gaussian model against these alternatives we study the following two cases; (i) the common factor has a non-Gaussian distribution, and (ii) the idiosyncratic factor has a non-Gaussian distribution. We estimate the parameters of each modelling choice through Maximum Likelihood. Since both non-Gaussian specifications include Gaussian as a special case, we use the likelihood ratio test to assess the fit of the unrestricted (non-Gaussian) over the restricted (Gaussian) model. Additionally, we assess the impact of these distributional assumptions on the capital requirements for ten portfolios of different loan types for the entire US Banking System. The results show that both non-Gaussian alternatives provide a better fit for case (i).

Moreover, the Skew Student's *t* specification does not provide a better fit over the Skew Normal. We quantify the impact of the non-Gaussian modelling choices on the capital charges and find that the degree of under(over) estimation depends on the sign and the absolute value of the shape parameter estimate of the Skew Normal.

To our knowledge this is the first empirical paper that develops a methodology and examines the impact of non-Gaussianity on the distribution of portfolio credit losses and on capital charges. Non-Gaussian process has not been studied before possibly due to the fact that it is technically challenging to implement, and for the case of the Skew Student's *t* distribution the estimation is computationally intensive. In particular, as explained later in Section 4, the estimation loss distribution involves the use of non-standard quadrature functions within the optimization routine.

The remainder of this paper is organized as follows. Section 2 briefly reviews Vasicek's original (1987, 2002) single factor model and the generalized version derived by Schönbucher (2001). Section 3 discusses the statistical properties of the Skew Normal and the Skew Student's *t* densities. Section 4 presents the estimation framework which is based on maximum likelihood. Section 5 describes the data sets. In Section 6, we present the estimation results and assess the impact of non-Gaussianity on capital requirement calculation. Finally, Section 7 provides some concluding remarks.

2. A Review of Vasicek (1987) and its Generalization

In this section, we first review the derivation of Vasicek's (1987, 2002) single factor limiting loss distribution and its underlying assumptions. Then, we describe the extensions to non-gaussian distributions made by Schönbucher (2001) which opens the way for our empirical specifications through *SN* and *ST* as we show later in the text.

Finally, we show how the capital requirements are computed under the generalized distribution.

The individual loss due to obligor *i* is defined as the product of the Exposure At Default (*EAD*), the Loss Given Default (*LGD*) and a default indicator variable (D_i) as follows:

$$L_i = EAD_i \times LGD_i \times D_i \tag{2.1}$$

The variable D_i is a Bernoulli random variable that takes the value one if the obligor defaults and zero otherwise. This setup implicitly assumes that EAD_i and LGD_i are time invariant for each obligor.[3] Then, the portfolio loss rate, L, can be calculated as

$$L = \frac{\sum_{i=1}^{n} L_i}{\sum_{i=1}^{n} EAD_i} = \sum_{i=1}^{n} w_i \times LGD_i \times D_i \tag{2.2}$$

where $w_i = \left(EAD_i \Big/ \sum_{i=1}^{n} EAD_i \right)$ is the portfolio weight for the *i*th loan. Vasicek (1987) assumes that the size of EAD_i and LGD_i are the same for all obligors, and, moreover, that the recovery rate is equal to zero such that $LGD_i = 1$. Further assumption that $EAD_i = EAD$ leads to $w_i = w = 1/n$, and a homogeneous portfolio with loss rate:

$$L = \frac{\sum_{i=1}^{N} D_i}{n} \tag{2.3}$$

Note that the *N* default random variables D_i have been treated as independent of each other. To allow for correlation among the default variables D_i in (2.3), let the asset return, R_i, for obligor *i* in the portfolio be driven by a single common factor *Y* and an idiosyncratic noise component ε_i:

$$R_i = \sqrt{\rho} Y + \sqrt{1-\rho}\,\varepsilon_i \tag{2.4}$$

where *Y* and ε_i are assumed to be mutually and serially independent random variables that follow a standardized Gaussian distribution $N(0,1)$, and $\sqrt{\rho}$ and $\sqrt{1-\rho}$ are the

[3] The *LGD* can be treated as a stochastic variable without changing the model results as long as it is assumed to be independent of D_i.

corresponding factor loadings. Note that under this specification, the asset returns of all firms are multivariate Normal with the same pairwise correlation ρ.

Vasicek (1987) assumes that the credit portfolio is fine grain, i.e. it consists of a large number of relatively small exposures. If this assumption holds, the idiosyncratic risk associated with the individual exposures will cancel out and only systematic risks that affect all the exposures will have an impact on the portfolio value and loss rate.

So far the default process has been treated as exogenous. Following Merton (1974), Vasicek (1987) assumes that the *i*th obligor defaults if the value of its assets, $A_{i,T}$, at loan maturity, falls below the debt contractual value, $B_{i,T}$. In the context of the credit portfolio model and assuming that all obligors have the same default probability, i.e. $pd_i = pd$, the endogenous default process based on Merton's (1974) can be introduced if D_i is defined as:

$$D_i = 1 \quad \text{if} \quad R_i \le \Phi^{-1}(pd) \quad \text{and} \quad D_i = 0 \quad \text{if} \quad R_i > \Phi^{-1}(pd) \qquad (2.5)$$

where $\Phi(\cdot)$ is the cumulative Gaussian distribution function and *pd* is the unconditional default probability. The default process defined in eq.(2.5) depends on the latent random variable *Y* that drives the asset returns R_i as follows:

$$\begin{aligned} p(y) &= P(D_i = 1 | Y = y) \\ &= P(R_i \le \Phi^{-1}(pd) | Y = y) \\ &= P\left(\sqrt{\rho} \cdot Y + \sqrt{1-\rho} \cdot \varepsilon_i \le \Phi^{-1}(pd) \middle| Y = y\right) \\ &= P\left(\varepsilon_i \le \frac{\Phi^{-1}(pd) - \sqrt{\rho} \cdot Y}{\sqrt{1-\rho}} \middle| Y = y\right) \\ &= \Phi\left(\frac{\Phi^{-1}(pd) - \sqrt{\rho} \cdot y}{\sqrt{1-\rho}}\right) \end{aligned} \qquad (2.6)$$

which is the probability of default conditional on the value of *Y*. Conditional on the realization *y* of *Y*, the individual defaults happen independently of each other. Thus, the unconditional probability of observing exactly *k* defaults is the average of the conditional probabilities of k defaults, averaged over the possible realizations of Y and weighted with the probability density function $\phi(y)$:

$$\begin{aligned} P\left[\sum_{i=1}^{n} D_i = k\right] &= \int_{-\infty}^{\infty} P\left[\sum_{i=1}^{n} D_i = k \middle| Y = y\right]\phi(y)\,dy \\ &= \int_{-\infty}^{\infty} \binom{n}{k}\left(p(y)\right)^k \left(1-p(y)\right)^{n-k} \phi(y)\,dy \end{aligned} \tag{2.7}$$

Vasicek (1987) showed that if the portfolio is large, then the law of large numbers ensures that the fraction of obligors that actually default is (almost surely) exactly equal to the individual default probability. From eq.(2.7), Vasicek shows that the limiting loss distribution for the homogeneous portfolio loss rate is:

$$F_L(l; pd, \rho) = \Pr[L \le l] = \Phi\left(\frac{\sqrt{1-\rho}\,\Phi^{-1}(l) - \Phi^{-1}(pd)}{\sqrt{\rho}}\right) \tag{2.8}$$

The loan portfolio loss distribution is fully determined by two parameters: the probability of default (*pd*) and the asset correlation (ρ). The former fixes the expected loss rate of the portfolio, while the latter controls the shape of the loss distribution. The density of $F_L(l; PD, \rho)$ can be derived by using the inverse function theorem and this is equal to:

$$f_L(l; pd, \rho) = \sqrt{\frac{1-\rho}{\rho}} \exp\left\{-\frac{\left(\sqrt{1-\rho}\,\Phi^{-1}(l) - \Phi^{-1}(pd)\right)^2}{2\rho} + \frac{\left(\Phi^{-1}(l)\right)^2}{2}\right\} \tag{2.9}$$

Schönbucher (2001) extends the single factor results of Vasicek to cases where the common and the idiosyncratic factors have non-Gaussian distributions. In particular, Schönbucher (2001) showed that if $Y \sim G(\cdot)$, and $\varepsilon_i \sim H(\cdot)$ for all *i*, and Y and ε_i (both centred and standardized) are independent, then the limiting loan loss distribution is equal to:

$$F_L(l; pd, \rho) = 1 - G\left(\frac{K}{\sqrt{\rho}} - \sqrt{\frac{1-\rho}{\rho}} H^{-1}(l)\right) \qquad (2.10)$$

where K is the default barrier which is given by the inverse of the asset return distribution. Notice that this extension adds flexibility and could potentially be very important for modelling real data. However, this extension comes at the cost of implementation complexities. This is because the default barrier K is no longer equal to the Gaussian inverse of *pd*. In particular, K is now equal to the inverse of the function that arises from the sum of the assumed distributions for *Y* and ε_i in eq.(2.4).

According to the advanced and foundation approach contained in Basel II, the capital requirements are computed as the difference between the unexpected loss (*UL*) and the expected loss (*EL*) scaled by the *LGD* and the effective remaining maturity. In this paper, we omit the effect of the time to maturity (see Kjersti (2005)). In the Basel II framework, banks are expected to cover their *EL* on an ongoing basis, because it represents just another cost component of the lending business. Therefore, under this methodology, capital is only needed for covering unexpected losses. Hence, banks are required to hold capital against *UL* and this corresponds to the CreditVaR of the portfolio. The capital requirements per unit of exposure are computed in this paper as:

$$\begin{aligned} CR_\alpha &= LGD \times (UL_\alpha - EL) \\ UL_\alpha &= H\left(\frac{K - \sqrt{\rho} G^{-1}(1-\alpha)}{\sqrt{1-\rho}}\right) \\ EL &= pd \end{aligned} \qquad (2.11)$$

where *UL* is obtained by inverting Schönbuchers (2001) generalized limiting loss distribution for a significance level α which, under Basel II, is set equal to 99.9%. Note that in eq.(2.11), the value of *LGD* depends on the type of loan under analysis and this value is prescribed in Basel II.

3. Specifying Non- Normality for the Generalized Vasicek Distribution

This section presents density specifications for functions $G(\cdot)$ and $H(\cdot)$ in equation (2.10) using the Skew Normal and the Skew Student's *t* densities. It provides a brief account of their main properties to facilitate the understanding of their application in the context of credit loss estimation.

3.1 The Skew Normal

The Skew Normal distribution proposed by Azzalini (1985, 1986) has density function:

$$f_z(z;\alpha) = 2\phi(z)\Phi(\alpha z), \qquad -\infty < z < \infty \tag{3.1}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative normal distribution functions, respectively, and α is the shape parameter with $-\infty < \alpha < \infty$. With the density function in eq.(3.1), we write $Z \sim SN(\alpha)$. In practice, we rarely work with this form of the skew normal density. If $Z \sim SN(\alpha)$ and $W = \mu + \sigma Z$, where $\mu \in (-\infty, \infty)$ is the location parameter and $\sigma \in (0, \infty)$ is the scale parameter, then we shall write $W \sim SN(\mu, \sigma^2, \alpha)$ and W has density:

$$f_W(w;\mu,\sigma,\alpha) = \frac{2}{\sigma}\phi\left(\tfrac{w-\mu}{\sigma}\right)\Phi\left(\alpha\cdot\left(\tfrac{w-\mu}{\sigma}\right)\right), \qquad -\infty < w < \infty, \tag{3.2}$$

The Skew Normal density has four special properties: (i) $SN(0)$ is $N(0,1)$; (ii) as $\alpha \to \infty$, $f_Z(z;\alpha)$ tends to the half-normal density; (iii) if *Z* is a $SN(\alpha)$ random variable, then –*Z* is a $SN(-\alpha)$; and (iv) $f_Z(z;\alpha)$ is strongly unimodal, i.e. $\log f_Z(z;\alpha)$ is a concave function of *z*.

Azzalini (1985) shows that the first four moments of the standardized Skewed Normal random variable *Z* are:

$$E[Z] = b\delta \tag{3.3}$$

$$\operatorname{var}[Z] = 1 - (b\delta)^2 \tag{3.4}$$

$$\gamma_1(Z) = \frac{(4-\pi)}{2} sign(\alpha) \left[\frac{\{E(Z)\}^2}{\operatorname{var}(Z)} \right]^{3/2} \tag{3.5}$$

$$\gamma_2(Z) = 2(\pi - 3) \left[\frac{\{E(Z)\}^2}{\operatorname{var}(Z)} \right]^{2} \tag{3.6}$$

where $b = \sqrt{\frac{2}{\pi}}$, $\delta = \alpha \big/ \sqrt{1+\alpha^2}$, $sign(\bullet)$ is a function that returns the sign of its argument and γ_1, γ_2 denote the third and fourth standardized cumulants. [4] Azzalini (1985) shows that, for the Skewed Normal distribution, the maximum value of skewness is about 0.995, while that for kurtosis is 0.869.

Figure 1 shows the effect of increasing the magnitude of the shape parameter value on the shape of the *SN* density. In Panel A, V_i are functions of $Z_i \sim SN(\alpha_i)$ such that: $V_i = (Z_i - \mu_{Z_i}) / \sigma_{Z_i}$. Therefore, it is possible to show that $V_i \sim SN\left(-\frac{\mu_{Z_i}}{\sigma_{Z_i}}, \frac{1}{\sigma_{Z_i}}, \alpha_i\right)$, where $\alpha_i \in \{0, 4, 10, +\infty\}$ for $i = 1, 2, 3, 4$ for the four cases. For increasing positive values of α, the Skew Normal density is right skewed, the mass of the density concentrates on the left and the right tail will always be heavier than that of the Normal. Conversely, Panel B shows the effect of increasing the magnitude of α when α is negative. The random variables $V_i^*, (i = 1,2,3,4)$ are exactly the same as V_i, but V_i^* are defined for the

4 The skewness is defined as the third standardized moment, while the kurtosis can be defined as the fourth standardized moment. Alternatively, the skewness and the kurtosis can be defined as the third and fourth standardized cumulants.

corresponding negative α values. In this case, the Skew Normal density is left skewed, the mass of the density concentrates on the right and the right tail will always be thinner than that of the Normal.

The Skew Normal is more flexible than the Normal because we can regulate the skewness and the excess kurtosis through the shape parameter α, albeit to a moderate degree. Since there is only one variable, we cannot regulate skewness and excess kurtosis at the same time. In this regards, the Skew Student's *t* is more flexible as skewness and excess kurtosis are separately controlled by two independent parameters.

3.2 Skew Student's *t* Distribution

The Skew Student's *t*-distribution[5], $ST(\alpha,v)$, has density function:

$$f_X(x;\alpha,v) = 2t_v(x)T_{v+1}\left(\alpha \cdot x\sqrt{\frac{v+1}{x^2+v}}\right), \qquad -\infty \le x \le \infty \tag{3.7}$$

where $t_v(\cdot)$ is the density of the standard Student's *t*-distribution with *v* degrees of freedom with $0 < v < \infty$; $T_{v+1}(\cdot)$ is the distribution function of the standard Student's *t*-distribution with $v+1$ degrees of freedom and α is the shape parameter with $-\infty < \alpha < \infty$. If $X \sim ST(\alpha,v)$ and $M = \mu + \sigma X$, where $\mu \in (-\infty,\infty)$ is the location parameter and $\sigma \in (0,\infty)$ is the scale parameter, then $M \sim ST(\mu,\sigma^2,\alpha,v)$ and M has density:

$$f_M(m;\mu,\sigma,\alpha,v) = 2\frac{1}{\sigma}t_v\left(\tfrac{m-\mu}{\sigma}\right)T_{v+1}\left(\alpha \cdot \left(\tfrac{m-\mu}{\sigma}\right)\sqrt{\frac{v+1}{\left(\tfrac{m-\mu}{\sigma}\right)^2+v}}\right), \qquad -\infty \le m \le \infty \tag{3.8}$$

The Skew Student's *t* density has six special properties: (i) $ST(0,v)$ is the standard Student's *t*-distribution $T_{v+1}(x)$; (ii) as $v \to \infty$, the Skewed Student's *t*- density

[5] Branco and Dey (2001) provide the original specification. The notations here follow those of Azzalini and Capitanio (2003). See Kjersti (2006) for a survey of different *ST* specifications.

converges to the skew-normal density; (iii) if $\alpha = 0$ and $v \to \infty$, then X is a $N(0,1)$ random variable; (iv) as $\alpha \to \infty$, the Skew Student's *t* tends to the folded-*t* distribution; (v) if X is a $ST(\alpha,v)$ random variable, then $-X$ is a $ST(-\alpha,v)$; (vi) if $v=1$, then the Skewed Student's *t*-density becomes a Skew-Cauchy density. Azzalini and Capitanio (2003) provide expressions for the standardized moments of a Skewed Student's *t* random variable *X*:

$$E[X] = c\delta, \quad for\ v > 1 \tag{3.9}$$

$$Var[X] = \frac{v^2}{(v-2)} - (c\delta)^2, \quad for\ v > 2 \tag{3.10}$$

$$\gamma_1(X) = \delta c\left(\frac{(3-\delta^2)}{v-3} - \frac{3v}{v-2} + 2(\delta c)^2\right)\left(\frac{v}{v-2} - (\delta c)^2\right)^{-\frac{3}{2}}, \quad for\ v > 3 \tag{3.11}$$

$$\gamma_2(X) = \left(\frac{3v^2}{(v-2)(v-4)} - \frac{4(\delta c)^2 v(3-\delta^2)}{v-3} + \frac{6(\delta c)^2 v}{v-2} - 3(\delta c)^4\right)\left(\frac{v}{v-2} - (\delta c)^2\right)^{-2} - 3, \quad for\ v > 4 \tag{3.12}$$

where $c = \sqrt{(v/\pi)}\,\Gamma((v-1)/2)/\Gamma(v/2)$ and $\delta = \alpha/\sqrt{1+\alpha^2}$. The appealing modelling advantage of the Skew Student's *t* is that it allows the tail thickness to be controlled separately via the degrees of freedom parameter, *v*. The most important theoretical difference between the *ST* and the *SN* is that the former has no restriction on the range of values for skewness and kurtosis (see Azzalini (2005, pp.180)).

Figure 2 shows the shape of four *ST* densities for the same shape parameter value (i.e. $\alpha = 9$) and for different degrees of freedom parameter values. The variable P_i is a function of the random variable $X_i \sim ST(\alpha, v_i)$, such that: $P_i = (X_i - \mu_{X_i})/\sigma_{X_i}$ for $i \in \{1,2,3,4\}$. It is possible to show that $P_i \sim ST(-\mu_{X_i}/\sigma_{X_i}, 1/\sigma_{X_i}, \alpha, v_i)$ where $v_i \in \{3,5,30,\infty\}$ for the four cases. As the degrees of freedom parameter increases, the *ST* converges to the *SN*. In fact, when v=30, the shape of the *ST* density approximates that of *SN*. Since the four densities are positively skewed, the right tail will always be heavier than that of Normal. Moreover, the tails of the *ST* will always be heavier than that of *SN* for any given α.

In summary, to model the non-Gaussianity of the common and the idiosyncratic factors density, it is useful if the density has the following three properties: (i) "strict inclusion"[6] of the normal density; (ii) mathematical tractability; and (iii) cover a wide range of skewness and kurtosis values. The Skew Normal density fulfils the first two requirements in having some tractability, and in capturing skewness and kurtosis through its shape parameter. The Skew Student's *t*-density fulfils the second and third properties and has a great control over the skewness and kurtosis range through the shape and the degrees of freedom parameters.

4. Maximum Likelihood Estimation

The parameters of Schönbucher's (2001) generalized loss density can be estimated using maximum likelihood. The estimation is based on observed portfolio default rates. Following Düllmann and Trapp (2004), we assume that the systematic and idiosyncratic risk factors have no autocorrelation. The generalized probability density for the observed default rates l_t is given by:

$$\begin{aligned} f_{L_t}\left(l_t;\boldsymbol{\theta}\right) &= \frac{\partial F_L\left(l_t;\boldsymbol{\theta}\right)}{\partial l_t} \\ &= \sqrt{\frac{1-\rho}{\rho}}\, g\left(\frac{K-\sqrt{1-\rho}H^{-1}\left(l_t\right)}{\sqrt{\rho}}\right)\left(\frac{1}{h\left(H^{-1}\left(l_t\right)\right)}\right) \end{aligned} \tag{3.13}$$

6 A density will strictly include the normal density if for a particular value of one or more of its parameters we obtain the normal specification. A density will not have the strict inclusion property if the normal density results as one or more of its parameters tend to the limit.

where $g(\cdot)$ is the density for the common factor and $g = \frac{\partial G(l_t)}{\partial l_t}$; $h(\cdot)$ is the density of the idiosyncratic factor and $h = \frac{\partial H(l_t)}{\partial l_t}$; $\boldsymbol{\theta}$ is the vector of parameters and K is the default barrier which is given by the inverse function of the asset return distribution as a function of $\boldsymbol{\theta}$.

The objective is to maximize the following constrained log-likelihood function:

$$\max_{\boldsymbol{\theta}} \ln L(\boldsymbol{\theta}; l_1, \ldots, l_T) = \sum_{t=1}^{T} \ln\left(f_L(\boldsymbol{\theta}; l_t)\right) \qquad (3.14)$$

The $\boldsymbol{\theta}$ set may contain additional parameters depending on the choice for $g(\cdot)$ and $h(\cdot)$. Düllmann and Trapp (2004) show that for the Vasicek (1987) loss density,[7] the value of the parameters $\left(pd^{ml}, \rho^{ml}\right)$ that maximize eq.(3.14) has a closed form solution. Therefore, the maximization problem can be solved analytically. Moreover, for this Gaussian case, Düllmann and Trapp (2004) derive a closed-form solution of the asymptotic Cramer-Rao lower bounds for the standard deviation of the estimators.

The main challenge in introducing non-Gaussianity is the computation of the default barrier, which is equal to the inverse function of the asset return distribution. Compared with the Skew Student's *t*, the Skew Normal case is a relatively manageable task. We collect our analytical results for the Skew Normal case in the following proposition.

Proposition:

For an asset return process of the form

$$R_i = \sqrt{\rho} Y + \sqrt{1-\rho}\,\varepsilon_i \qquad (3.15)$$

(a) if $Y \sim SN(\alpha)$ and $\varepsilon_i \sim N(0,1)$, then

[7] Vasicek (1987) assumes that $g(.)$ and $h(.)$ are both standard normal densities.

$$R_i \sim SN\left(\frac{\sqrt{\rho}\cdot\alpha}{\sqrt{1+\alpha^2(1-\rho)}}\right) \tag{3.16}$$

(b) if $Y \sim N(0,1)$ and $\varepsilon_i \sim SN(\alpha)$, then

$$R_i \sim SN\left(\frac{\sqrt{1-\rho}\cdot\alpha}{\sqrt{1+\alpha^2\rho}}\right) \tag{3.17}$$

Proof: Using the method of moment generating functions we can show, for any real number *a* and *b*, that the proposal stated by Azzalini (2005) is

$$\frac{aU+bZ}{\sqrt{a^2+b^2}} \sim SN\left(\frac{b\alpha}{\sqrt{a^2\left(1+\alpha^2\right)+b^2}}\right) \tag{3.18}$$

where $U \sim N(0,1)$, $Z \sim SN(\alpha)$, and *U* and *Z* are mutually independent. Then, defining $a=\sqrt{1-\rho}$, $b=\sqrt{\rho}$, we obtain result in eq.(3.16). Also, defining $a=\sqrt{\rho}$, $b=\sqrt{1-\rho}$, we obtain result in eq.(3.17).□

To approximate the value of the default barrier *K*, i.e. the *pd*-quantile of the asset return distribution given in eq.(3.16) and eq.(3.17), we use the Cornish Fisher Expansion (see Cornish and Fischer (1960)).

For the Skew Student's *t* case, the default barrier does complicate the computation considerably. This is because the distribution followed by the sum of a Gaussian and a Skew Student's *t* is not known. Here, we compute the distribution via numerical quadrature (for a review on this topic, see Gander and Gautschi (2000) or Moler (2004)). Then, one can compute the quantile by minimizing the distance between the approximated distribution at a given probability level, $F_R(k)$ and the corresponding *pd* value.

We analyze the case where $Y \sim ST(\alpha, v)$ and $\varepsilon_i \sim N(0,1)$ to illustrate how this approach works. First, rewrite the SFM as $R_i = C + D_i$, where $C = \sqrt{\rho} \cdot Y$ and $D_i = \sqrt{1-\rho} \cdot \varepsilon_i$. Next, given that $Y \sim ST(0,1,\alpha,v)$, we have $C \sim ST(0,\rho,\alpha,v)$ and $D_i \sim N(0,1-\rho)$ for all *i*. Third, since R_i is the sum of two independent random variables, then the convolution $f_C * f_{D_i}$ of f_c and f_{D_i} is the function given by:

$$
\begin{aligned}
f_{R_i}(r_i) &= f_C(c) * f_{D_i}(d_i) \\
&= f_C(c) * f_{D_i}(r_i - c) \\
&= \int_{-\infty}^{\infty} f_C(c) f_{D_i}(r_i - c) dc \\
&= \int_{-\infty}^{\infty} \frac{2}{\sqrt{\rho}} t_v\left(\frac{c}{\sqrt{\rho}}\right) T_{v+1}\left(\alpha\left(\frac{c}{\sqrt{\rho}}\right)\sqrt{\frac{v+1}{\left(c/\sqrt{\rho}\right)^2 + \sqrt{\rho}}}\right)\left(\frac{1}{\sqrt{2\pi(1-\rho)}} e^{-\frac{1}{2}\left(\frac{r_i - c}{\sqrt{1-\rho}}\right)^2}\right) dc
\end{aligned}
\tag{3.19}
$$

Note that once we have integrated eq.(3.19), we got rid of *c*, and the remaining expression for the density is a function of r_i. Finally, the distribution function of R_i can be obtained by integrating eq.(3.19) wrt r_i as follows:

$$
\begin{aligned}
F_{R_i}(k) &= \Pr\left[R_i \le k\right] = \int_{-\infty}^{k} f_{R_i}(r_i) dr_i \\
&= \int_{-\infty}^{k}\int_{-\infty}^{\infty} f_C(c) \times f_D(r_i - c) dc dr_i
\end{aligned}
\tag{3.20}
$$

Now, combine the two steps, where

$$
F_{R_i}(k) = \int_{-\infty}^{k}\int_{-\infty}^{\infty} \frac{2}{\sqrt{\rho}} t_v\left(\frac{c}{\sqrt{\rho}}\right) T_{v+1}\left(\alpha\left(\frac{c}{\sqrt{\rho}}\right)\sqrt{\frac{v+1}{\left(c/\sqrt{\rho}\right)^2 + \sqrt{\rho}}}\right)\left(\frac{1}{\sqrt{2\pi(1-\rho)}} e^{-\frac{1}{2}\left(\frac{r_i - c}{\sqrt{1-\rho}}\right)^2}\right) dc dr_i \tag{3.21}
$$

We use an adaptive Simpson quadrature to numerically evaluate the double integral. The Simpson quadrature was computed such that it approximates the integral to within an error of 10^{-8}. To compute the *pd*-quantile, we simply solve the following nonlinear problem wrt *k*:

$$F_{R_i}(k) - pd = 0 \tag{3.22}$$

The tolerance level was set equal to $1e^{-8}$. All computations were performed in Matlab.

5. The Federal Reserve Aggregate Loss Data

Our data sample consists of quarterly sector aggregate charge-off rates (not seasonally adjusted) for all US commercial banks starting from Q1:1985 to Q3:2007. The charge-off rates are published by the Federal Reserve Board on a quarterly basis. The charge-off rates for any category of loan are defined as the flow of a bank's net charge-offs (gross charge-offs minus recoveries) during a quarter divided by the average level of its loans outstanding in that quarter.[8] The charge-off series is reported at three aggregate levels. At the top level, we have the 'Commercial Banking System' which consists of 'Business', 'Consumer', 'Loans Secured by Real Estate', 'Agricultural' and 'Leases'. The 'Consumer loans', in turn, consists of 'Credit Cards' and 'Other Consumer loans', while the 'Loans Secured by Real Estate' consists of 'Mortgages'[9] and 'Commercial Real Estate'.[10] Since the charge-off rates are net of recoveries which can be from any period in the past, charge-off rates can sometime have negative or zero values.[11] This happens whenever the recovery amount for a quarter is greater than or equal to gross charge-off of that quarter. We replace all non-positive charge-off rates by the series

---

8 As published, these ratios are multiplied by 4x100 to convert to annual percentage rates.

9 Mortgage loans include loans secured by one-to four-family properties, including home equity lines of credit.

10 Commercial real estate loans include construction and land development loans, loans secured by multifamily residences, and loans secured by nonfarm, non-residential real estate. The data for the Mortgage and Commercial Real Estate are available from Q1:1991.

11 There are other problems associated with the use of banks' charge-off rates. As Lamb and Perraudin (2006) noted, when a new manager takes over a division of a bank, he or she may wish to write off delinquent and semi-delinquent loans in order to be able to demonstrate a better performance subsequently. Nevertheless, following Lamb and Perraudin's (2006), we assume that the aggregation of many banks charge-offs will help to remove any possible bias due to such actions.

minimum positive historical value. Following Lamb and Perraudin (2006), we scaled the series by $(1/LGD)$. This is because the charge-off rates are published net of recoveries. The respective LGD for each loan portfolio was taken from Basel Committee on Banking Supervision (2004).

The time series plots of the ten charge-off rate series are shown in Figure 3. Panel A shows the relationship between the scaled charge-off rate for the 'Banking System' and its main sectors; Panel B and C show the 'Consumer' and 'Real Estate' sectors with their respective sub-components while Panel D shows the non-scaled 'Banking System' and its constituent sectors. In Table 1 we present some descriptive statistics for the scaled (by $(1/LGD)$) and non-scaled charge-off rate series.

Given the scale of the sub-prime financial crisis, some readers might be surprised by the relatively low level of the real estate charge-off series. The observed levels are relatively low because the sub-prime crisis is largerly related to financial instruments (i.e. CDO's) that do not form part of the bank's balance sheet. Moreover, it is not likely to detect signs of deterioration in the banking system from our historical data set since it will take some time before the bad loans are charge-off from the system.

From the time series plot of the charge-off rate in Figure 3, Panel A, one can observe that the 'Banking System' series remains stable even though there is a clear deterioration in 'Business' and 'Lease' that starts in 1999. Moreover, the 'Consumer' series exhibits high levels of charge off compared to all the other series. Since only 'Banking System' and 'Loans Secured by Real Estate' share the same time series properties but not the others, it suggests that 'Loans Secured by Real Estate' must be the largest component of 'Banking System'. It is important to note that the "Banking

System” series is the only that was not scaled by the *LGD*. The unscaled series is presented in Panel D.

Table 1 reports in its first two rows the unconditional mean and standard deviations of the unscaled loss rates. This table is useful for comparing the entire ‘Banking System’ against its subcomponents. The table reveals that ‘Credit Cards’ has the highest loss rate at 2.16% which is more than doubles that of the ‘Banking System’ (0.84%). By contrast, ‘Mortgages’ exhibit a small loss rate of just 0.15%.

‘Agricultural’ loans have the highest volatility (1.06%), measured as the standard deviation of the series, and this represents approximately three times that of the ‘Banking System’ (0.36%). This is not too surprising given the high levels of loss rate in the ‘Agricultural’ sector at the beginning of the sample period. Volatility of the other series, eg ‘Credit Cards’ (1.02%), ‘Commercial Real Estate’ (0.71%), ‘Consumer’ (0.58%) and ‘Business’ (0.56%), is relatively high compared with the entire ‘Banking System’. ‘Mortgages’ has the lowest volatility (0.07%), much smaller than the volatility of the other series. Overall, the statistic suggests that ‘Mortgages’ is less risky, and ‘Credit Cards’ Loans are most risky. Nevertheless, as remarked by Lamb and Perraudin (2006), a more important aspect of the riskiness of a loan type is its asset correlation within the sector. This and other sources of risk will be analyzed in the next section.

Since we estimate the model for the scaled series it is important to study the skewness and kurtosis of the scaled series that are reported in the bottom panel of Table 1. Note that the mean and standard deviation for the scaled ‘Consumer’ series, the ‘Credit Card loan’ in particular, are much higher than the other series. This is partly because the *LGD* set by Basel II for these sectors is also higher.

6. Empirical Results on US Credit Portfolio Losses

In this section we report the parameter estimates for Vasicek's (1987) Gaussian case for the ten sets of observed charge-off rates. Then, we show the results for the case where the common factor follows a Skew Normal or a Skew Student's *t*-distribution. We assess the fit of these non-Gaussian models to the observed charge-off rates and compare them against that of the Gaussian case. Moreover, we also compare the fit of the *ST* against that of the *SN*, and compare the impact of these two alternative non-Gaussian specifications on capital charges. Finally, we repeat the analysis for the case where the idiosyncratic factor is either *SN* or *ST* distributed.

6.1 The Vasicek's (1987) Gaussian Model

The estimation results of the base case Vasicek model with Gaussian common and idiosyncratic risk factors are presented in Table 2. This model assumes that the *pd*, the probability of default, and ρ, the correlation between the asset returns of any pair of firms, are constant for all firms and across all time periods.

The *pd* parameter is an estimator of the expected charge-off rate. Therefore, its value is close to the mean reported in the bottom panel of Table 1. Credit cards (6.5%), Consumer Loans (3.32%), Business Loans (1.88%) and Other Consumer (1.56%) have the highest estimated default probabilities.

The correlation parameter ρ determines the shape of the loss density, and consequently, its quantiles. The square root of ρ measures the correlation between the asset return and the single common factor. The higher the ρ, the stronger is the sector's exposure to fluctuations in the common factor which is believed to be driven by the business cycle. According to Table 2, ρ for Commercial Real Estate Loans (28.37%), Agricultural Loans (16.61%), Loans Secured by Real Estate (11.11%), Business (8.31%) and Leases (6.74%) are among the highest suggesting that these sectors are the most sensitive to changes in the economic conditions. All these portfolios have a higher ρ when compared to that of the 'Banking System' (2.07%). Note that 'Business' is the only portfolio that appears in the high *pd* and high ρ group, whereas none of the three consumption series has high ρ's. This result is not too surprising given that the consumer portfolios typically represented by a large number of small heterogenous loans, whereas the 'Business' portfolios tends to be dominated by a smaller number of large loans.

Under some regularity conditions (see Greene (2000, pp.127)), the maximum likelihood estimator follows an asymptotic Gaussian distribution. The asymptotic standard deviation of the ML parameters can be estimated with: (i) the inverse of the Hessian; (ii) the outer product of gradients (OPG), which is also known as the Berndt, Hall, Hall and Haussman estimator; and (iii) the Sandwich or Quasi-Maximum-Likelihood Estimator (see White (1982)). All three estimates are computed and reported in Table 2 which shows that all parameters estimates are significant at the 1% level regardless of the choice of standard error estimator. We also drew 1,000 bootstrap data samples for each of the charge-off series. This is because the *pd* and ρ parameters are constrained to the interval [0,1] and this might cause the asymptotic distribution of the estimates not to be Normal. However, as shown by Table 2, the bootstrap estimator is of the same magnitude as the asymptotic estimators.

Table 2, Panel B reports the Jarque-Bera normality test[12] if the sampling distribution of *pd* and ρ follows a standard normal distribution. For the case of the *pd*, the Jarque-Bera test rejected normality for the case of 'Agricultural' and 'Commercial Real Estate'. Regarding ρ, the Jarque Bera rejected normality for the case of the 'Banking System', 'Business', 'Credit Cards' and 'Mortgages Loans'. We can conclude that it is not clear that the MLE will be normally distributed for all series. Given that the standard error for the parameter estimates is almost identical between the bootstrap and the QMLE estimator, we will report only the QMLE standard errors.

6.2 Non-Gaussian Common Factor

Tables 3 and 4 report the results, respectively, for the cases where the common factor follows a *SN* and a *ST* distribution while the idiosyncratic factor follows the standard Gaussian distribution. The *pd* and ρ estimates for the Gaussian Vasicek base case reported in Table 2 are repeated here for ease of comparison. From Table 3 one can see that: (i) all *pd* and ρ parameters are statistically significant at the 1% level; (ii) the *pd* estimates are very similar to the Vasicek's base case; (iii) the ρ parameters increase with respect to the Vasicek's base case, the increase is greater whenever there is a significant shape parameter α; (iv) the shape parameter is not statistically significant for three portfolios, viz. 'Credit Card', 'Other Consumer' and 'Commercial Real Estate', and this suggests that the *SN* specification does not provide a significantly better fit than that of the Gaussian case for these three portfolios; (v) negative shape parameters were observed for 'Loans Secured by Real Estate' (-9.5), 'Mortgages' (-7.6), 'Banking System' (-3.2) and 'Agricultural Loans' (-2.9); (vi) the corresponding skewness

[12] We also performed Lilliefors normality tests for *pd* and ρ, but we omitted the results given that these were similar to the Jarque Bera.

coefficients for the sectors listed in (v), as shown in Table 3, are -0.95, -0.92, -0.70 and -0.7, respectively.

Since the *SN* distribution contains the normal as a particular case, we can assess the fit of $Y \sim SN(\alpha)$ against that of $Y \sim N(0,1)$ with the log-likelihood ratio test (*LR*). Table 3 shows that the *LR* test is significant at the 10% level for 7 out of ten series ('Banking System', 'Business', 'Consumer', 'Loans Secured by Real Estate', 'Mortgages', 'Agricultural' and 'Leases'). Note that the three series for which the *SN* does not provide a better fit are also the ones that did not have a statistically significant shape parameter.

Figure 4 plots the distribution of charge-off rates fitted under Vasicek's Gaussian density and that of the $SN(\alpha)$ common factor against the historical observed rates. It is clear that the $SN(\alpha)$ provides a marked improvement in the fit especially in cases where $|\alpha|$ is large.

Regarding the *ST* results shown in Table 4, we provide the following observations: (i) the magnitudes of the *pd*, ρ and α parameters are almost identical to those for the *SN* case, except for 'Agricultural loans'; (ii) the same three portfolios that did not have a significant shape parameter in the *SN* case remain insignificant under the *ST* (viz. 'Credit Card', 'Other Consumer', 'Commercial Real Estate'); (iii) the relationship reported in point (iii) of the *SN* case between α and ρ also holds here for the *ST* ; (iv) the estimated degrees of freedom parameter v is very large in all sectors except for the agricultural series.

Regarding the fit of the *ST*, we can clearly see from the *LR* listed at the bottom of Table 4 that: (i) the *ST* provides a better fit than Vasicek's Gaussian alternative for exactly the same sectors where *SN* also provided a better fit; (ii) the *ST* does not provide a statistically better fit than the *SN*. This result corresponds to the large estimated values for v and the fact that the *ST* collapses to *SN* as $v \to \infty$. The only exception to point (ii) is the 'Agricultural Loan' portfolio, where the estimated v is 7.3033.

6.3 Non-Gaussian Idiosyncratic Noise

The results in Tables 5 and 6 show the parameter estimates for the case where the idiosyncratic factor is *SN* and *ST* distributed, respectively, while the common factor is normally distributed. The log-likelihood ratio shows that in none of the ten series, the non-Gaussian alternative for the idiosyncratic factor provides a better fit than the Normal distribution.

6.4 Impact on Capital Charge

Table 7 compares the capital charge per unit of exposure for the cases where the common factor is Gaussian, *SN* or *ST*. Note that the capital charges for the *SN* and the *ST* cases are higher for negative shape parameter values (viz. 'Commercial Banking System', 'Loans Secured by Real Estate', 'Mortgages' and 'Agricultural Loans'), and lower for positive shape parameter (viz. 'Business loan', Consumer loan', 'Credit Card', 'Other Consumer loan', 'Commercial Real Estate' and 'Lease'). The difference in capital charge estimates becomes smaller the closer the shape parameter is to zero. Note that the skewness has a large impact on the capital charges. For example, the capital requirement for 'Real Estate' is more than double under *SN* than under Gaussian. It is clear that higher negative skewness in asset returns leads to higher capital requirements. For the case of the *SN*, the estimated skewness is close to the maximum admissible skewness of *SN*, which corresponds to roughly double the capital requirements. It is important to recall that as the degrees of freedom parameter tends to infinite, then the *ST* converges to the *SN*. Thus, lower estimates for the degrees of freedom parameter lead to higher capital requirements. The only portfolio with a small degrees of freedom parameter estimate is 'Agricultural Loans'. In this particular case, the estimated degrees of freedom parameter is 7.3 (see Table 4) and this explains the difference between the capital requirement for the *ST* common factor (0.09) and that of the *SN* (0.05).

The last two panels of Table 7 show that the capital requirements under the *SN* and the *ST* idiosyncratic noise assumption are similar and do not differ significantly form those of the Vasicek Gaussian assumption. This result might be due to the fact that within each portfolio, the idiosyncratic risk is well diversified, and there is no significant exposure to idiosyncratic risk associated with individual exposures. This is plausible since our data represent large portfolios. Nevertheless, this result should be taken with care. If we were to repeat this exercise on the loan portfolio of small or medium sized commercial banks, then the result and the conclusion might change.

7. Concluding Remarks

Vasicek (1987, 2002) derive a limiting loan portfolio loss distribution which is founded on a stochastic asset return process that is driven by a common and an idiosyncratic factor both of which are Gaussian. Schönbucher (2001) extends the Vasicek model to include cases where the common and the idiosyncratic factors are non-Gaussian. Despite its analytical tractability and the rich theoretical insights which have heavily influenced Basel II, the literature lacks direct empirical support.

In this paper, we have developed a methodology to model empirically the impact of non-Gaussian risk factors on credit loss distributions and capital charge. We allow the underlying common and idiosyncratic factors to be Skewed Normal and Skewed Student's *t* individually. The main interest of using non-Gaussian distributions is to control for the effect of the asset return skewness and excess kurtosis on the shape of the loss distriburion. The maximum likelihood of our models require further analytical results of functions of non-Normal variables which was then performed to official charge-off rates published by the Federal Reserve Board for ten U.S. sector charge-off rates.

The main finding of our paper is that non-Gaussian modelling provides a significantly improved fit in the loss density for seven out of the ten portfolios analyzed. The most conclusive finding is that the common factor should be best modelled as Skewed Normal. Allowing the common factor to be Skewed Student's *t* or the idiosyncratic factor to be non-Gaussian, does not provide noticeably significant improvement to the empirical fit.

Our findings confirm that non-Gaussian modelling of the common factor is very important, and highlight the inadequacy of the existing Basel II framework. The capital requirements obtained by assuming a Gaussian distribution for the asset return could over-or underestimate the capital requirements. This degree of over-or underestimation depends on the sign and the magnitude of the skew parameter of the Skew Normal. Large negative skew parameter value leads to an underestimated capital requirement, while large positive skew parameter leads to an overestimation.

The non-Gaussian modelling of the idiosyncratic factor did not produce any insignificant impact possibly because the ten sectors analyzed here are large portfolios, and the idiosyncratic risks might have already been cancelled out. Due diligence should be observed when the loan portfolio is small or not well diversified.

Our empirical evidence suggests that Skew Normal is an adequate representation of the distributional properties of the latent common factor, since the estimated degrees of freedom for the Skew Student's *t*-distribution takes very high values, approaching a Skew Normal. However, we propose using Skew Student's *t* as a modelling choice because this distribution adds extra flexibility and has the potential to accommodate both heavy tails and skewness, which might prove useful for modelling the losses due to the credit crisis when the data becomes available.

References

[1] Azzalini, A., (1985). “A class of distributions which includes the Normal ones.” *Scandinavian Journal of Statistics* 12, pp.171–178.

[2] Azzalini, A., (1986). “Further results on a class of distributions which includes the Normal ones.” *Statistica* XLVI, pp. 199–208.

[3] Azzalini, A. and A. Capitanio (2003). “Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution.” *Journal of the Royal Statistical Society B* 65, pp.579–602.

[4] Azzalini, A. (2005). “The skew-Normal distribution and related multivariate families.” *Scandinavian Journal of Statistics* 32, pp.159–188.

[5] Basel Committee on Banking Supervision (2004). International convergence of capital measurement and capital standards. A revised framework. Consultative document.

[6] Batiz, E., Christodoulakis, G., and Poon, S., (2008). “Credit loss distributions: Vasicek and beyond.” Working Paper.

[7] Branco, M. D. and Dey, D.K. (2001). “A general class of multivariate skew-elliptical distributions.” *Journal of Multivariate Analysis* 79, pp.99–113.

[8] Cornish, E. A. and Fisher, R. A. (1960). *Technometrics*, 2, pp.209–225.

[9] Düllmann, K., and Trapp, M. (2004). “Systematic risk in recovery rates- An empirical analysis of U.S. corporate credit exposures.” Deutsche Bundesbank, pp.1–35.

[10] Gander, W. and Gautschi, W. (2000). “Adaptive quadrature-revisited”. *BIT Numerical Mathematics* 40, pp.84–101. Available at www.inf.ethz.ch/personal/gander

[11] Greene, W. (2000). Econometric Analysis. Upper saddle river, New Jersey, Prentice-Hall, Inc.

[12] Kjersti, A. (2005). “The Basel II IRB approach for credit portfolios: A survey.” Norsk Regnesentral (Norwegian Computing Center, NR); available at www.nr.no.

[13] Kjersti, A. (2006). “The generalized hyperbolic skew student’s *t*-distribution.” *Journal of Financial Econometrics* 4, pp. 275–309.

[14] Merton, R. (1974). “On the pricing of corporate debt: The risk structure of interest rates.” *Journal of Finance* 29, 449–470.

[15] Moler, C. (2004). Numerical computing with Matlab. Philadelphia, USA, SIAM (Society for Industrial and Applied Mathematics). Available at www.mathworks.com/moler

[16] Perraudin, W. and Lamb, R. (2006). “Dynamic loan loss distribution: estimation and implications.” Working paper. Imperial College, Tanaka Business School, London.

[17] Schönbucher, P. (2001). “Factor models: Portfolio credit risks when defaults are correlated.” *The Journal of Risk Finance*, 3, 45–56.

[18] Vasicek, O. (1987). “Probability of loss on loan portfolio.” KMV Corp.; available at www.kmv.com.

[19] Vasicek, O. (2002). “Loan portfolio value.” *Risk* 15, 160–162.

[20] White, H. (1982). “Maximum likelihood estimation of misspecified models.” *Econometrica*, 53, pp.1–16.

Table 1. Descriptive statistics for sector charge-off rates (not seasonally adjusted)
Published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
Charge-Off Series[a]										
Mean	0.0084	0.0084	0.0216	0.0423	0.0102	0.0034	0.0044	0.0015	0.0064	0.0049
Std. Dev.	0.0036	0.0056	0.0058	0.0102	0.0033	0.0034	0.0071	0.0007	0.0106	0.0030
Scaled Charge-Off Series by (1/*LGD*)										
LGD (from Basel II)	1[b]	0.45	0.65	0.65	0.65	0.35	0.35	0.35	0.45	0.45
Mean	0.0084	0.0186	0.0332	0.0650	0.0156	0.0097	0.0125	0.0042	0.0143	0.0110
Std. Dev	0.0036	0.0124	0.0089	0.0156	0.0051	0.0096	0.0202	0.0020	0.0236	0.0067
Skewness	1.0391	0.6689	0.1049	0.4605	0.6348	1.4497	1.9341	1.8200	2.9656	0.7795
Kurtosis	3.5180	2.6329	2.3442	3.1237	3.2945	4.4513	5.6632	7.2608	12.0876	3.6823
No of observations	91	91	91	91	91	91	67†	67†	91	91

(a) The sector charge-off rates (not seasonally adjusted) are defined as the flow of a bank's net charge-offs (gross charge-offs minus recoveries) during a quarter divided by the average level of loans outstanding in that quarter.

(b) The 'Commercial Banking System' comprises 'Business', 'Consumer', 'Secured by Real Estate', 'Agricultural' and 'Lease' Loans. The share of losses for the five sectors is not disclosed by the Fed and Basel II does not provide any value for the *LGD* of 'Commercial Banking System'. For simplicity, we assume that the *LGD* for the 'Commercial Banking System' is equal to 1.

(†) The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.

Table 2. Parameter estimates for Vasicek loss distribution of sector charge-off rates (not seasonally adjusted)
Published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
PANEL A										
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0122	0.0042	0.0128	0.0113
Inv. Hessian Std Error	0.0004	0.0015	0.0010	0.0016	0.0005	0.0010	0.0027	0.0002	0.0017	0.0009
QMLE Std Error	0.0004	0.0014	0.0009	0.0016	0.0005	0.0011	0.0028	0.0002	0.0022	0.0007
OPG Std Error	0.0003	0.0018	0.0011	0.0016	0.0005	0.0010	0.0027	0.0002	0.0014	0.0011
Bootstrap Std Error	0.0004	0.0013	0.0009	0.0016	0.0005	0.0010	0.0028	0.0002	0.0022	0.0007
ρ	0.0207	0.0831	0.0155	0.0149	0.0166	0.1111	0.2837	0.0190	0.1661	0.0674
Inv. Hessian Std Error	0.0030	0.0113	0.0023	0.0022	0.0024	0.0146	0.0351	0.0032	0.0205	0.0093
QMLE Std Error	0.0025	0.0078	0.0021	0.0020	0.0022	0.0102	0.0342	0.0034	0.0242	0.0104
OPG Std Error	0.0038	0.0167	0.0026	0.0024	0.0027	0.0225	0.0362	0.0035	0.0209	0.0093
Bootstrap Std Error	0.0025	0.0079	0.0020	0.0020	0.0022	0.0101	0.0348	0.0033	0.0237	0.0103
Log-Likelihood	396.3	282.4	299.8	251.8	355.7	335.7	247.6	336.2	309.1	331.3
PANEL B:										
ρ JBera Statistic	2.81	3.91	33.09	22.06	7.80	2.07	1.24	17.50	0.26	1.95
ρ Critical Value	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93
PD JBera Statistic	1.20	1.47	1.57	0.97	2.04	4.26	29.53	5.33	53.27	1.68
PD Critical Value	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93	5.93

† The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.

The Jarque-Bera test is a two sided goodness of fit test suitable when a fully-specified null distribution is unknown and its parameters must be estimated. The test statistic is $JB = (n/6)\left(s^2 + (k-3)^2/4\right)$ where n is the sample size, s is the sample skewness, and k is the sample kurtosis. For large sample sizes, the test statistic has a chi-square distribution.

In Matlab, the Jarque-Bera test uses a table of critical values computed using Monte-Carlo simulation for sample sizes less than 2000 and significance levels between 0.001 and 0.50. Critical values for a test are computed by interpolating into the table, using the analytic chi-square approximation only when extrapolating for larger sample sizes.

Table 3. Parameter estimates for Vasicek loss distribution where $Y \sim SN(\alpha)$ and $\varepsilon_j \sim N(0,1)$ on sector charge-off Rates (not seasonally adjusted) published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
Base case results from Table 3:										
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0122	0.0042	0.0128	0.0113
ρ	0.0207	0.0831	0.0155	0.0149	0.0166	0.1111	0.2837	0.0190	0.1661	0.0674
PD	0.0084	0.0191	0.0333	0.0650	0.0156	0.0104	0.0122	0.0042	0.0137	0.0111
QMLE Std Error	0.0004	0.0012	0.0009	0.0016	0.0005	0.0013	0.0028	0.0002	0.0025	0.0007
ρ	0.0496	0.2007	0.0377	0.0155	0.0215	0.2722	0.2837	0.0522	0.3074	0.1564
QMLE Std Error	0.0078	0.0245	0.0067	0.0142	0.0124	0.0251	0.0342	0.0096	0.0464	0.0266
α	-3.2535	4.3759	3.2299	0.2664	0.7597	-9.5118	0.0176	-7.5864	-2.9389	4.1673
QMLE Std Error	1.0840	1.7034	1.2925	2.9647	1.1602	3.7864	0.0198	3.8098	0.6606	2.3698
Log-Likelihood	398.8	284.1	301.9	251.8	355.7	341.9	247.6	341.3	315.0	336.0
Log-Lik Vasicek	396.3	282.4	299.8	251.8	355.7	335.7	247.6	336.2	309.1	331.3
LR ratio	4.9	3.5	4.2	0.0	0.1	12.4	0.0	10.1	11.8	9.4
p-value	0.0264*	0.0607+	0.0404*	0.9812	0.8100	0.0004#	0.9999	0.0015#	0.0006#	0.0021#
CF Skewness	-0.7037	0.8137	0.7005	0.0040	0.0718	-0.9515	0.0000	-0.9279	-0.6573	0.7982
CF Exc. Kurtosis	0.5475	0.6645	0.5442	0.0005	0.0261	0.8186	0.0000	0.7917	0.4999	0.6477

Notes:
† The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.
CF refers to the Common Factor
'#' indicates cases where the skew normal provides a significant better fit than the normal at the 1% level.
'*' indicates cases where the skew normal provides a significant better fit than the normal at the 5% level.
'+' indicates cases where the skew normal provides a significant better fit than the normal at the 10% level.
Given that the results by using different estimators are similar, we report only the inference based on the QMLE standard error.

Table 4. Parameter estimates for portfolio loss distribution where $Y \sim Skew-t\left(0,1,\alpha,DF\right)$ and $\varepsilon_j \sim N(0,1)$ on sector charge-off rates (not seasonally adjusted) published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
Vasicek Gaussian base case results from Table 2:										
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0122	0.0042	0.0128	0.0113
ρ	0.0207	0.0831	0.0155	0.0149	0.0166	0.1111	0.2837	0.0190	0.1661	0.0674
Results for the Skew Normal common factor from Table 3:										
PD	0.0084	0.0191	0.0333	0.0650	0.0156	0.0104	0.0122	0.0042	0.0137	0.0111
ρ	0.0496	0.2007	0.0377	0.0155	0.0215	0.2722	0.2837	0.0522	0.3074	0.1564
α	-3.2535	4.3759	3.2299	0.2664	0.7597	-9.5118	0.0176	-7.5864	-2.9389	4.1673
PD	0.00844	0.0191	0.0333	0.0650	0.0156	0.0104	0.0131	0.0042	0.0145	0.0111
QMLE Std Error	0.00037	0.0012	0.0009	0.0016	0.0005	0.0013	0.0029	0.0002	0.0027	0.0006
ρ	0.04957	0.2006	0.0376	0.0150	0.0214	0.2721	0.3547	0.0522	0.2150	0.1496
QMLE Std Error	0.00774	0.0202	0.0062	0.0020	0.0125	0.0231	0.0828	0.0096	0.1283	0.0268
α	-3.25007	4.3744	3.2278	0.1127	0.7550	-9.5100	-1.0195	-7.5848	-2.0343	4.1390
QMLE Std Error	1.06705	1.0742	1.1360	0.0339	1.1745	0.5069	0.9314	3.8522	1.2285	2.5810
DF	1607	3694	2270	2689	1010	4092	33.5455	1995	7.3033	43.6796
QMLE Std Error	21.8575	5.2737	635.74	14.4795	188.9755	97.9466	0.9967	32.6412	7.7082	0.1940
Log-Likelihood	398.8	284.1	301.9	251.8	355.7	341.9	247.7	341.3	315.3	336.0
Log-Lik Vasicek	396.3	282.4	299.8	251.8	355.7	335.7	247.6	336.2	309.1	331.3
LR ratio	4.9	3.5	4.2	0.0	0.01	12.4	0.2	10.1	12.5	9.5
p-value	0.0263*	0.0602+	0.0403*	0.9445	0.8091	0.0004#	0.6979	0.0015#	0.0003#	0.0021#
Log-Lik Y~SN(α)	398.8	284.1	301.9	251.8	355.7	341.9	247.6	341.3	315.0	336.0
LR ratio	0.014	0.0106	0.0087	0.0054	0.0137	0.0089	0.1507	0.0063	0.6728	0.0493
p-value	0.9058	0.9181	0.9259	0.9414	0.9067	0.9247	0.6979	0.9367	0.4121	0.8243

Notes
'#', '*' and '+' indicate cases where the skew normal provides a significant better fit than the normal at the 1% level, 5% level and 10% level, respectively.
† The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.

Table 5. Parameter estimates for Vasicek loss distribution where $Y \sim N(0,1)$ and $\varepsilon_j \sim SN(\alpha)$ on sector charge-off rates (not seasonally adjusted) published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
Base case results from Table 3:										
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0122	0.0042	0.0128	0.0113
ρ	0.0207	0.0831	0.0155	0.0149	0.0166	0.1111	0.2837	0.0190	0.1661	0.0674
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0124	0.0042	0.0127	0.0113
QMLE Std Error	0.0004	0.0014	0.0009	0.0016	0.0005	0.0011	0.0027	0.0002	0.0022	0.0007
ρ	0.0177	0.0826	0.0155	0.0069	0.0139	0.0967	0.2571	0.0190	0.1427	0.0674
QMLE Std Error	0.0022	0.0077	0.0021	0.0461	0.0019	0.0090	0.0373	0.0034	0.0209	0.0104
α	-1.000	0.095	0.000	1.350	-1.850	-1.148	-0.865	0.014	-1.929	-0.012
QMLE Std Error	0.0090	0.0360	0.0000	9.2810	0.1780	0.2200	0.5220	0.0040	0.2670	0.0020
Log-Likelihood	396.4	282.4	299.8	251.8	355.7	335.8	247.7	336.3	309.97	331.3
Log-Lik Vasicek	396.3	282.4	299.8	251.8	355.7	335.7	247.6	336.2	309.1	331.3
LR ratio	0.2044	0.0003	0.0000	0.0077	0.0599	0.3312	0.0751	0.0000	1.7522	0.0000
p-value	0.6512	0.9857	0.9996	0.9299	0.8065	0.5649	0.7840	0.9991	0.1854	0.9988

† The Fed started reporting the charge-off rate for the ‘Commercial Real Estate’ and ‘Mortgages’ only from Q1:1994.

Table 6. Parameter estimates for Vasicek loss distribution where $Y \sim N(0,1)$ and $\varepsilon_j \sim Skew-t(0,1,\alpha,DF)$ on sector charge-off rates (not seasonally adjusted) published by the Federal Reserve Board for the period Q1:1985 to Q3:2007

	Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
<u>Base case results from Table 3:</u>										
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0122	0.0042	0.0128	0.0113
ρ	0.0207	0.0831	0.0155	0.0149	0.0166	0.1111	0.2837	0.0190	0.1661	0.0674
PD	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0124	0.0042	0.0117	0.0113
QMLE Std Error	0.0001	0.0001	0.0009	0.0025	0.0009	0.0015	0.0034	0.0001	0.0011	0.0001
ρ	0.0156	0.0846	0.0167	0.0078	0.0123	0.0969	0.2578	0.0195	0.1589	0.0567
QMLE Std Error	0.0022	0.0077	0.0021	0.0461	0.0019	0.0090	0.0373	0.0034	0.0209	0.0104
α	-1.011	0.0961	0.0000	1.3708	-1.898	-1.1476	-0.8698	0.01434	-1.9225	-0.0131
QMLE Std Error	0.0012	0.0261	0.0000	10.2810	0.1780	0.2200	0.5220	0.0040	0.2670	0.0020
DF	4240	4836	8950	4230	4678	3862	4835	6110	2398	6330
QMLE Std Error	43.9	10.9	998.3	32.6	201.8	118.3	3.9	64.3	65.3	38.2
Log-Likelihood	396.4	282.4	299.8	251.8	355.7	335.8	247.7	336.3	309.9	331.3
Log-Lik Vasicek	396.3	282.4	299.8	251.8	355.7	335.7	247.6	336.2	309.1	331.3
LR ratio	0.2045	0.0003	0.0000	0.0078	0.0602	0.3322	0.0796	0.0000	1.7755	0.0000
p-value	0.6511	0.9862	0.9982	0.9296	0.8062	0.5644	0.7778	0.9992	0.1827	0.9984

† The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.

Table 7. Capital charges for Vasicek's (1987) single factor Gaussian model and the non-Gaussian alternatives

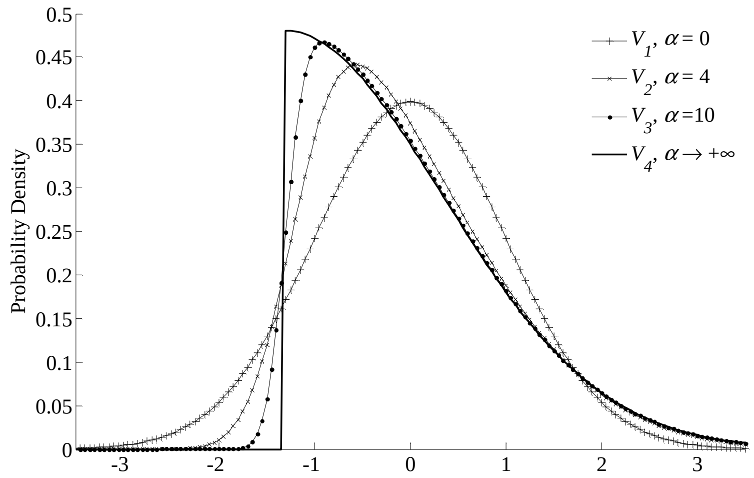
		Commercial Banking System	Business Loans	Consumer Loans	Credit Card	Other Consumer Loans	Loans Secured by Real Estate	Commercial Real Estate Loans†	Mortgages†	Agricultural Loans	Lease
	LGD	1	0.45	0.65	0.65	0.65	0.35	0.35	0.35	0.45	0.45
	Sig.	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990
Vasicek	EL	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0124	0.0042	0.0127	0.0113
	UL	0.0247	0.1073	0.0718	0.1260	0.0383	0.0819	0.2376	0.0129	0.1431	0.0628
	CR	0.0162	0.0398	0.0251	0.0396	0.0147	0.0253	0.0789	0.0030	0.0587	0.0232
SN Common factor	shape	-3.2535	4.3759	3.2299	0.2664	0.7597	-9.5118	0.0176	-7.5864	-2.9389	4.1673
	EL	0.0084	0.0191	0.0333	0.0650	0.0156	0.0104	0.0122	0.0042	0.0137	0.0111
	UL	0.0329	0.0630	0.0588	0.1259	0.0374	0.1657	0.2376	0.0189	0.2273	0.0354
	CR	0.0245	0.0198	0.0165	0.0396	0.0142	0.0543	0.0789	0.0051	0.0961	0.0109
ST Common factor	shape	-3.2501	4.3744	3.2278	0.1127	0.7550	-9.5100	-1.0195	-7.5848	-2.0343	4.1390
	EL	0.0084	0.0191	0.0333	0.0650	0.0156	0.0104	0.0131	0.0042	0.0145	0.0111
	UL	0.0329	0.0630	0.0588	0.1260	0.0374	0.1656	0.2577	0.0189	0.1263	0.0348
	CR	0.0245	0.0198	0.0165	0.0396	0.0142	0.0543	0.0856	0.0051	0.0503	0.0107
SN Idiosyncratic Factor	shape	-1.0000	0.0950	0.0000	1.3500	-1.8500	-1.1480	-0.8650	0.0140	-1.9290	-0.0120
	EL	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0124	0.0042	0.0127	0.0113
	UL	0.0249	0.1073	0.0718	0.1253	0.0387	0.0857	0.2561	0.0129	0.1527	0.0628
	CR	0.0165	0.0398	0.0251	0.0392	0.0150	0.0266	0.0853	0.0030	0.0630	0.0232
ST Idiosyncratic Factor	shape	-1.011	0.0961	0.0000	1.3708	-1.898	-1.1476	-0.8698	0.01434	-1.9225	-0.0131
	EL	0.0084	0.0188	0.0332	0.0650	0.0156	0.0096	0.0124	0.0042	0.0127	0.0113
	UL	0.0249	0.1073	0.0718	0.1253	0.0387	0.0857	0.2561	0.0129	0.1527	0.0628
	CR	0.0165	0.0398	0.0251	0.0392	0.0150	0.0266	0.0853	0.0030	0.0630	0.0232

Notes in Table 7: shape=Estimated Shape Parameter; EL=Expected Loss, UL=Unexpected Loss, Cap.Req=Capital Requirements, SN=Skew Normal, ST=Skew Student's *t*

Consistent with Basel II, the confidence level to compute the capital requirements is set to 99.9%. This means that an institution is expected to suffer losses that exceed its economic capital once in a thousand years on average.† The Fed started reporting the charge-off rate for the 'Commercial Real Estate' and 'Mortgages' only from Q1:1994.

Figure 1. Skew Normal Density

PANEL A: Skew Normal density function for positive shape values



PANEL B: Skew Normal density function for negative shape values

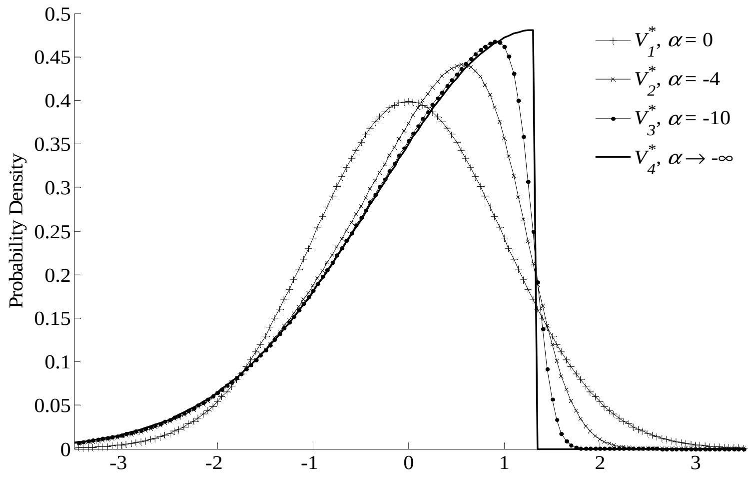


Table of moments for Panel A

Random Variable	V_1	V_2	V_3	V_4
Shape (α)	0	4	10	+∞
Delta (δ)	0	0.9701	0.9950	1
Expected Value	0	0	0	0
Variance	1	1	1	1
Std Dev	1	1	1	1
Skewness $\gamma_1(V_i)$	0	0.7844	0.9556	0.9953
Excess Kurtosis $\gamma_2(V_i)$	0	0.6328	0.8232	0.8692

Table of moments for Panel B

Random Variable	V_1^*	V_2^*	V_3^*	V_4^*
Shape (α)	0	-4	-10	-∞
Delta (δ)	0	-0.9701	-0.9950	1
Expected Value	0	0	0	0
Variance	1	1	1	1
Std Dev	1	1	1	1
Skewness $\gamma_1(V_i)$	0	-0.7844	-0.9556	-0.9953
Excess Kurtosis $\gamma_2(V_i)$	0	0.6328	0.8232	0.8692

Figure 2. Skew Normal and Skew Student's *t* density function for positive shape values

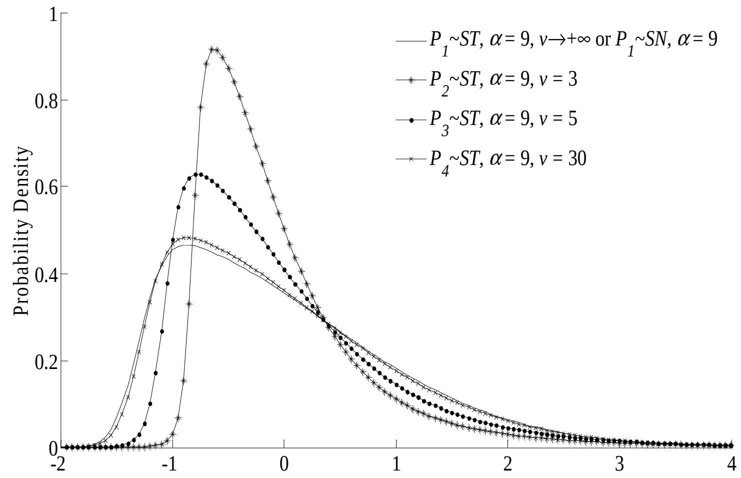


Table of moments for Figure 2

Random Variable	P_1	P_2	P_3	P_4
Shape (α)	9	9	9	9
Degrees of freedom (v)	$+\infty$	3	5	30
Delta (δ)	0.9939	0.9939	0.9939	0.9939
Expected Value	0	0	0	0
Variance	1	1	1	1
Std Dev	1	1	1	1
Skewness $\gamma_1(P_i)$	0.9556	nd	2.5029	1.0803
Excess Kurtosis $\gamma_2(P_i)$	0.8232	nd	19.6611	1.3676

nd = not defined

Figure 3. US Federal reserve board quarterly annualized charge-off rates for the period Q1:1985 to Q3:2007

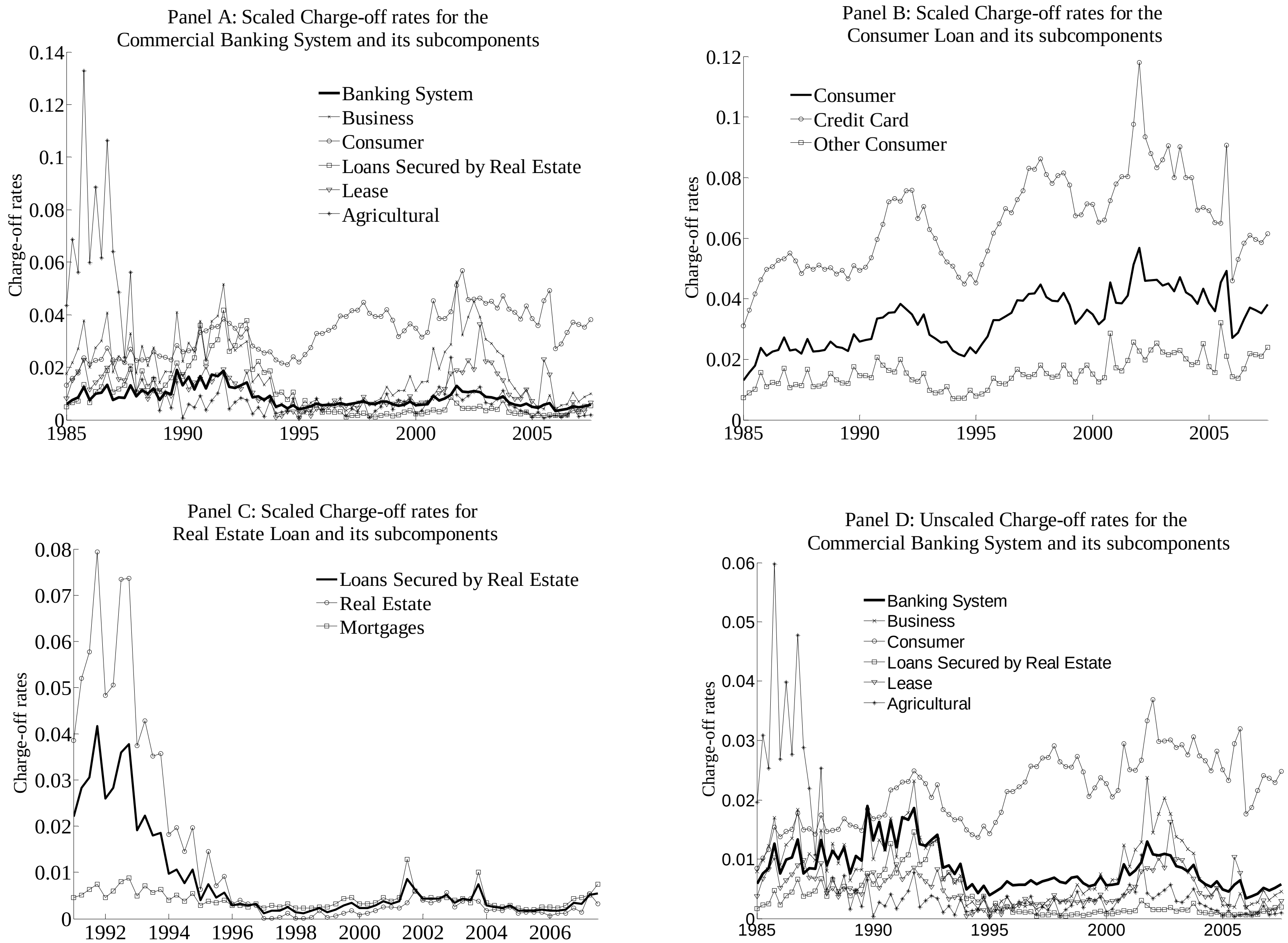


Figure 4. Distribution of quarterly charge-off rate under Normal and Skew Normal common factors against the historical distribution

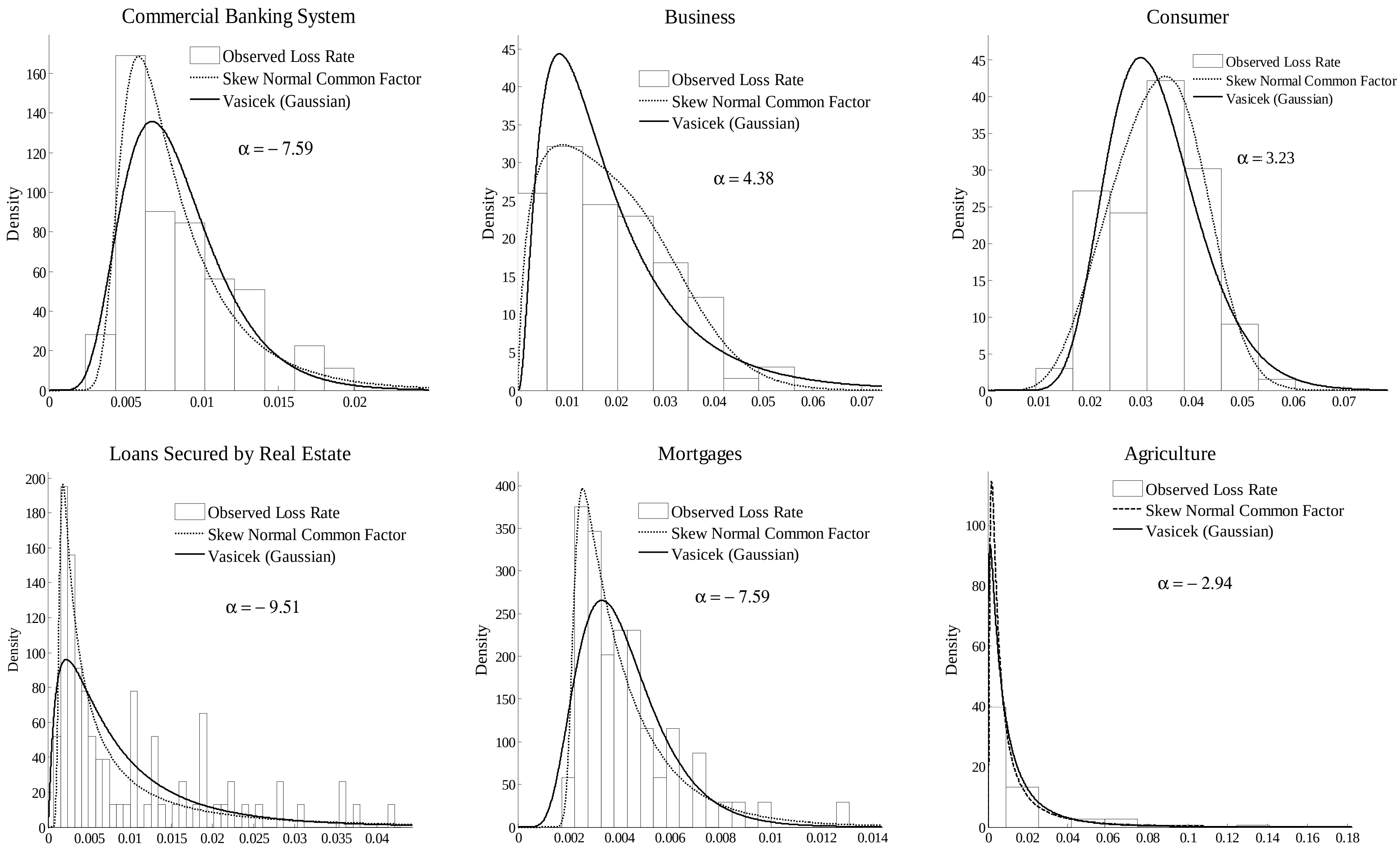


Figure 4 (Continued)

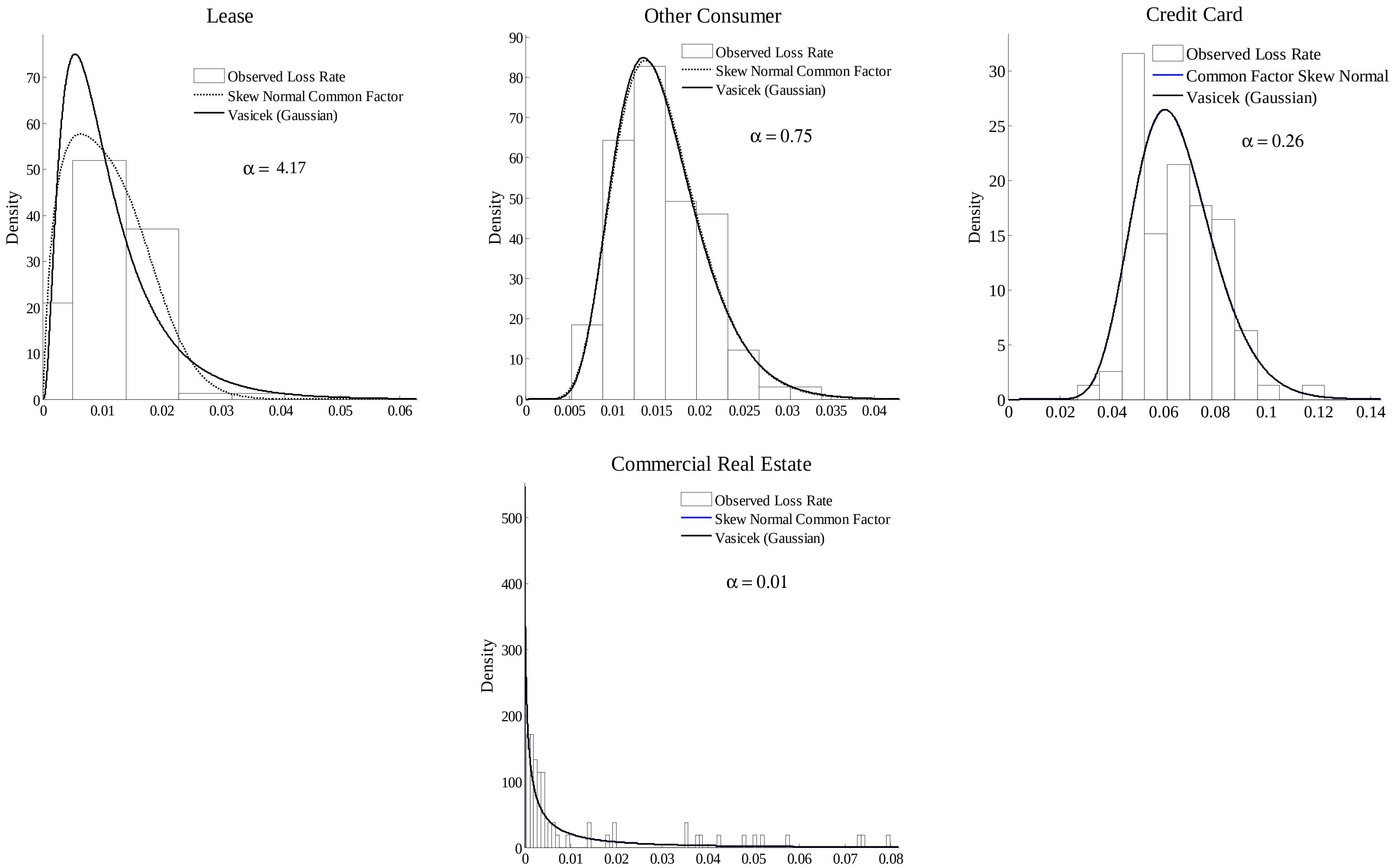


Figure 4 (Continued)

Vasicek's loss density assumes that the asset return process is driven by a single Gaussian common factor. Vasicek's loss density depends on the value of two parameters: the probability of default (*pd*) and the asset correlation (ρ). The values for these parameters for each portfolio are taken from Table 2. The Skew Normal common factor modelling is an alternative to Vasicek's Gaussian proposition that provides a superior fit for seven out of ten analyzed portfolios. The non-Gaussian density depends on three parameters: (i) the probability of default (*pd*); (ii) the asset correlation coefficient (ρ); and (iii) the shape parameter (α). The values of these parameters for each portfolio are taken from Table 3.