Institute of Economic Studies, Faculty of Social Sciences
Charles University in Prague

Loss, Default, and Loss Given Default Modeling

Jiří Witzany

IES Working Paper: 9/2009

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Witzany, J. (2009). " Loss, Default, and Loss Given Default Modeling " IES Working Paper 9/2009. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Loss, Default, and Loss Given Default Modeling

Jiří Witzany*

*University of Economics, Prague
E-mail: jiri.witzany@vse.cz

February 2009

Abstract:
The goal of the Basle II regulatory formula is to model the unexpected loss on a loan portfolio. The regulatory formula is based on an asymptotic portfolio unexpected default rate estimation that is multiplied by an estimate of the loss given default parameter. This simplification leads to a surprising phenomenon when the resulting regulatory capital depends on the definition of default that plays the role of a frontier between the unexpected default rate estimate and the LGD parameter whose unexpected development is not modeled at all or only partially. We study the phenomenon in the context of single-factor models where default and loss given default are driven by one systemic factor and by one or more idiosyncratic factors. In this theoretical framework we propose and analyze a relatively simple remedy of the problem requiring that the LGD parameter be estimated as a quantile on the required probability level.

Keywords: credit risk, correlation, recovery rate, regulatory capital

JEL: G21, G28, C14

Acknowledgements:
The research has been supported by the Czech Science Foundation grant no. 402/06/0890 “Credit Risk and Credit Derivatives”
Working paper, first version November 2008, this version February 2009.

1. Introduction

The Basle II regulatory formula (see Basle, 2006) aims to provide a sufficiently robust estimate of unexpected losses on banking credit exposures that should be covered by the capital. It is a compromise between the most advanced mathematical modeling techniques and the demand for a practical implementation. One of the most important simplifications is the decision to calculate unexpected losses (UL) using an estimate of the Unexpected Default Rate (UDR) multiplied through by the expected Loss Given Default parameter (LGD), i.e. UL=UDR·LGD. The capital requirement (C) as a percentage out of the exposure is then set equal to the difference between the unexpected and expected loss (EL), C = UL-EL = (UDR-PD)·LGD, where PD is the expected default rate, i.e. the probability of default.

While the expected default rate estimation based on the Vasicek (1987) approach is considered to be relatively robust, the resulting estimation of the unexpected loss has been criticized for neglecting the unexpected LGD (or equivalently recovery) risk. It has been empirically shown in a series of papers by Altman et al. (see e.g. 2004), Gupton et al. (2000), Frye (2000b, 2003), or Acharya et al. (2007) that there is not only a significant systemic variation of recovery rates but moreover a negative correlation between frequencies of default and recovery rates, or equivalently a positive correlation between frequencies of default and losses given default. Consequently the regulatory formula significantly underestimates the unexpected loss on the targeted confidence probability level (99.9%) and the time horizon (one year). Some authors have proposed alternative unexpected loss formulas incorporating the impact recovery risk variation.

Frye (2000a, 2000b) has used a single systemic factor model with an idiosyncratic factor driving the event of default and another independent idiosyncratic factor driving the recovery rate. The loading of the systemic factor for modeling of default and recovery rates may differ. The recovery rate is modeled as a normal variable truncated at 100%. Frye does not provide an analytical formula but analyzes robustness of the loss estimates using Monte Carlo simulation for different combinations of the input parameters. The parameters are also estimated using the maximum likelihood method from the Moody's Risk Service Default database. Alternatively Dullmann and Trapp (2004) apply the logit transformation for recovery modeling in the same set up as Frye.

Pykhtin (2003) considers a single systemic factor model where default is driven by a systemic factor and an idiosyncratic factor while recovery is driven not only by the systemic factor and an independent idiosyncratic factor, but at the same time by another idiosyncratic factor driving the obligor's default. The collateral (recovery) value is set to have a lognormal distribution. Pykhtin arrives to an analytic formula which requires numerical approximations of the bivariate normal cumulative distribution values. The author admits that calibration of the model is difficult.

Tasche (2004) proposes a single factor approach modeling directly the loss function. If there is no default the value of the loss function is zero and if there is a default (the systemic factor exceeds the default threshold) the value of the loss is drawn from a distribution as a function the systemic factor. The obligor factor is decomposed as usual into the systemic and idiosyncratic factor. In other words the single obligor factor is used to model the event of default and the loss given default as well. Tasche proposes to model LGD by a beta distribution. Quantiles of the loss function conditional on the systemic factor values may be expressed as an integral over a tail of the normally distributed factor. Tasche proposes to approximate the integral using Gauss quadrature and tests the model for different PD, mean/variance LGD, and correlation values. The approach is also elaborated in Kim (2006).

This study is motivated not only by the fact that the Basle II formula significantly underestimates the unexpected credit losses but also by the observation according to which the regulatory capital requirement depends on the definition of default which in a sense puts a border line between the PD and LGD parameters. This phenomenon has been analyzed in Witzany (2008) using a Merton model based simulation. To give a more tractable analytical explanation we will apply the Tasche and Frye single factor models as benchmarks against which we analyze the sensitivity of the regulatory formula. At the same time we propose a simple specification of the regulatory formula in order to eliminate the problem. We propose to preserve the formula $UL = UDR \cdot LGD$ as well as the regulatory formula for unexpected default rate (UDR), but to reinterpret the parameter LGD as the 99.9% quantile of possible portfolio loss given default values. The Basle (2005) document goes in this direction requiring LGD estimates to incorporate potential economic downturn conditions and adverse dependencies between default rates and recovery rates but fails to specify the confidence probability level of those conservative estimations. We argue that any probability level below 99.9% preserves the problem definition of default sensitivity (and underestimation of the 99.9% loss function percentile) while the 99.9% LGD quantile solves the problem under reasonable modeling assumptions. We propose a single factor beta distribution based technique calibrated with account level LGD mean, variance, cure rate and a correlation to obtain robust estimates of the 99.9% LGD quantiles. As the reinterpretation of the formula leads to significantly higher capital requirement we propose to reduce the probability level e.g. to a more a realistic 99.5% currently used by the Solvency II proposal.

2. Sensitivity of the Regulatory Capital on the Definition of Default

According to Basle II the contribution of a receivable to the unexpected loss of a well-diversified portfolio as a percentage of the exposure is estimated by the formula

$$UL = UDR \cdot LGD, \text{ where}$$

$$UDR = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{r} \cdot \Phi^{-1}(0{,}999)}{\sqrt{1-r}}\right). \tag{1}$$

The correlation r is set up by the regulator (e.g. 15% for mortgage loans, 4% for revolving loans, and somewhere between the two values depending on PD for other retail loans) while the parameters PD and LGD are estimated by the bank (in the IRBA approach).

The usual LGD estimation approach is based on a sufficiently large historical data set of a homogenous portfolio of receivables A in terms of product type, credit rating, and collateralization. The receivables have been observed for at least one year and we have a record $l : A \to [0,1]$ of percentage losses $l(a)$ on the exposure at default if default occurred or 0 otherwise for every $a \in A$, and an indicator function $d : A \to \{0,1\}$ of default in the one year horizon. It seems natural to require that $d(a) = 1$ iff $l(a) > 0$ as in Tasche (2004), however in practice such a condition is difficult to achieve. According to Basle II receivables more than 90 days overdue must be marked as defaulted. Some of the clients then naturally happen to pay all their obligations back; in particular in case of retail clients days overdue may just be a result of payment indiscipline not of a real lack of income to repay the loan. Hence we may require only that $l(a) > 0$ implies $d(a) = 1$ but not vice versa. The PD and LGD parameters of the Basle II formula (1) can be simply estimated from the given reference data set as

$$\bar{p} = \frac{|D|}{n},$$

$$\overline{lgd} = \frac{\sum_{a \in A} l(a)}{|D|} = \frac{\bar{l}}{\bar{p}},$$

where $D = \{a \in A \mid d(a) = 1\}$, $n = |A|$, and $\bar{l} = \frac{1}{n}\sum_{a \in A} l(a)$.

Here we are using equally weighted average loss given default that could be applied to a portfolio homogenous in terms of size. Let $D_H = \{a \in A \mid l(a) > 0\}$ be the set of receivables where we observed a positive loss, i.e. a hard default, and $\overline{p_H}$, $\overline{lgd_H}$ the averages as above. While the average (or expected) account level percentage loss $\bar{l} = \bar{p} \cdot \overline{lgd} = \overline{p_H} \cdot \overline{lgd_H}$ remains unchanged it is easy to see that $\overline{p_H} < \bar{p}$ and $\overline{lgd_H} > \overline{lgd}$ provided $D_H \subset D$. As banks have certain freedom for setting up their own definition of default, the ratio $\bar{p} / \overline{p_H}$ may be in practice anywhere between 1 and 2. Banks may choose a lower days-past-due default threshold (e.g. 60 days), or lower materiality condition (minimum amount past due implying the default), or apply different cross-default rules (default on one product implying defaults on other products with the same obligor), etc. More accounts with ultimate loss zero are then marked as defaulted. On the other hand the definition of default must not be too soft: if an account is marked as defaulted the probability of real loss should be at least 50%. Hence given the same historical information (reference data set A) with the account level average loss $EL = \bar{l}$ and choosing different definition of default we obtain different values of $PD \in \left(\overline{p_H}, 2 \cdot \overline{p_H}\right)$ and $LGD = EL / PD \in \left(\overline{lgd_H}, 2 \cdot \overline{lgd_H}\right)$. Since the definition of default does not change the distribution of losses implied by the reference data set the unexpected loss estimate given by (1) should remain essential the same. However Figure shows that this is not the case. When we set $EL = 2\%$ and let $PD \in (2.5\%, 5\%)$ then the $UL = UL(PD)$ parameter goes from 16.3% down to 12.5%. In other words choosing the softest possible definition of default will reduce the capital requirement $C = UL - EL$ by almost 30% compared to the hard definition of default.

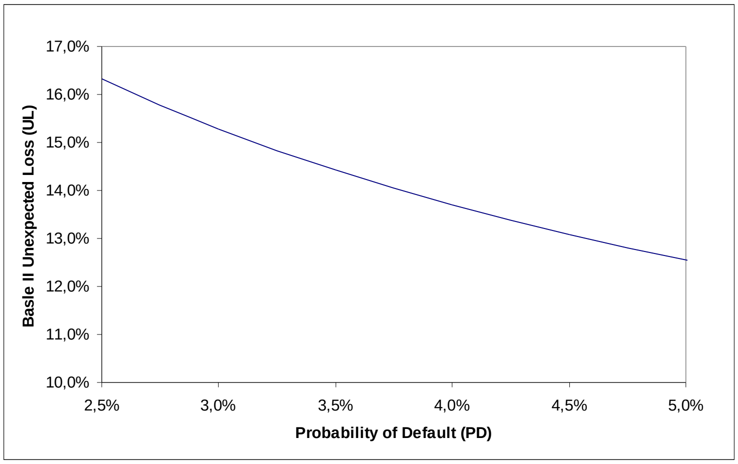


Figure 1. Unexpected loss according to the Basle II formula if EL=PD·LGD=2% is fixed and PD varies from 2.5% to 5% ($r = 15\%$).

It could be argued that the problem is solved by the requirement (Basle, 2005) on LGD to reflect downturn economic conditions or PD/LGD correlation. However this requirement given sufficiently rich historical data set is normally implemented using only the data set $A' \subset A$ from years with economic downturn conditions and/or high-observed frequency of default. The PD, LGD parameters estimated from A' and UL calculated according to (1) will again depend on the definition of default in the same way as above.

3. Alternative Single Factor Models

The single factor models of Frye (2000a, 2000b), Pykhtin (2003), Tasche (2004), and others can be generally described as follows. Let the (percentage) loss of a receivable in the given time horizon be an increasing function of one systemic factor X and of a vector $\vec{z}$ of idiosyncratic factors $L = L(X, \vec{z})$. The factor X captures macroeconomic or other systemic influences that may develop in time while $\vec{z}$ reflects specificities of each individual obligor in a portfolio. Hence the impact of $\vec{z}$ is diversified away in a large (asymptotic) portfolio while X remains as a risk factor. Consequently the future unknown loss on a large portfolio can be modeled as $E[L \mid X]$ (see Gordy, 2003 for details). Since we assume that L is increasing in X the problem to find quantiles of $E[L \mid X]$ reduces to calculation of quantiles of X. If x is the desired (e.g. for 99.9%) quantile of X then $UL = E[L \mid X = x]$. This is a clear advantage of the single-factor approach compared to the multi-factor approach where we work with a vector $\vec{X}$ of systemic factors instead of one factor X and the determination of quantiles of $E[L \mid \vec{X}]$ becomes complex.

The expression for the unexpected loss may be decomposed into two parts corresponding to the unexpected default rate and loss given default:

$E[L \mid X = x] = P[L > 0 \mid X = x] \cdot E[L \mid L > 0, X = x]$.

Here we use the hard definition of default $D_H \Leftrightarrow L > 0$ while as explained above in practice we usually need to work with a softer definition of default. We will say that $D = D(X, \vec{z})$ is a consistent notion of default provided $L > 0 \Rightarrow D$. Then the unexpected loss may be in general decomposed as

(2) $UL = P[D \mid X = x] \cdot E[L \mid D, X = x]$.

The simplest version of the singled-factor model is probably the model proposed by the Tasche (2004). The loss function $L = L(X, z)$ is driven by one standard-normally distributed factor $Y = \sqrt{r}X + \sqrt{1-r}z$ where X and z are independent standard-normally distributed, and r is their correlation. If L is assumed to have a cumulative probability distribution function $F_L : [0,1] \to [0,1]$ then we may express the loss function in the form $L(X, z) = F_L^*(\Phi(\sqrt{r}X + \sqrt{1-r}z))$ or just $L(Y) = F_L^*(\Phi(Y))$ where $F_L^*(z) = \inf\{t : F_L(t) \ge z\}$ is the generalized inverse of F_L.

In a sense more natural model has been proposed by Frye (2000a,2000b) which may be in a generalized for described as follows. Let $Y_1 = \sqrt{r_1}X + \sqrt{1-r_1}z_1$ and $Y_2 = \sqrt{r_2}X + \sqrt{1-r_2}z_2$ be two standard-normally distributed factors with one systemic and two independent idiosyncratic factors. The correlations r_1 and r_2 may be in general different. The first factor Y_1 drives defaults in the model while the second Y_2 is assumed to drive losses in case of default. I.e. there is a default threshold y_D and a nonnegative non-decreasing function G so that the loss function can be expresses as follows:

(3) $L(X, z_1, z_2) = \begin{cases} 0 \text{ if } Y_1 \le y_D, \\ G(Y_2) \text{ otherwise.} \end{cases}$

If F_G is the distribution function of the random variable $G(Y_1)$ then the loss function may be expressed as $G(Y_2) = F_G^*(\Phi(Y_2))$.

The Pykhtin (2003) model in a sense unifies the two models. In a generalized form let Y_1 be the driver of default as above, on the other hand let

$$(4)\qquad Y_2 = \sqrt{r_1}X + \sqrt{1-r_1}(\sqrt{w}z_1 + \sqrt{1-w}z_2)$$

be the driver of loss given default incorporating not only the systemic factor and a new idiosyncratic factor but also the idiosyncratic factor from Y_1. The loss function $L(X, z_1, z_2)$ is expressed by (3) as in the Frye's model. The approach enables us to model the fact that loss in case of obligors' default is determined by the value of assets and the specific financial situation at the time of default as well as by the workout/bankruptcy specific development.

Since we are in particular interested in unexpected loss given default modeling let us compare the three models in this respect. The unexpected loss given (hard) default conditional on a value of the systemic factor can is expressed as

$$(5)\qquad E[L \mid L>0, X=x] = \frac{1}{1-\Phi\left(\frac{y-\sqrt{r}x}{\sqrt{1-r}}\right)} \int_{\frac{y-\sqrt{r}x}{\sqrt{1-r}}}^{\infty} L(\sqrt{r}x + \sqrt{1-r}z) f(z) dz$$

where $y = \Phi^{-1}(P[L=0])$. On the other for the Frye model we get a nicer formula

$$(6)\qquad E[L \mid L>0, X=x] = E[G(Y_2) \mid X=x] = \int_{-\infty}^{+\infty} G(\sqrt{r_2}x + \sqrt{1-r_2}z) f(z) dz$$

since the value of $G(Y_2)$ does not depend on the idiosyncratic factor driving the default conditional on $X = x$. Regarding the Pykhtin model the reader may write down the double integral for $E[L \mid L>0, X=x]$ as an exercise. The approach if properly calibrated presents economically more faithful model compared to the other two. In fact the three correlation parameters of the model may be linked to the default correlation, loss given default correlation, and the default – loss given default correlation. Nevertheless since the model is difficult to calibrate and as is computationally complex we will on the Tasche and Frye models. In fact the models are two special cases of the Pykhtin model: set in (4) $w=1$ and $r_1 = r_2$ for the Tasche model and $w = 0$ for the Frye model.

It follows from the analysis done in the following Section (see e.g. Figure 3) that the density function in (5) concentrates significantly to the border of default. Consequently given a correlation $r = r_1 = r_2$ it is not surprising that the variance of $E[L \mid L>0, X]$ in the Tasche model given by (5) is much lower than the variance in the Frye's model given by (6). To model losses in case of default we will use the beta distribution with minimum 0 and maximum 1 determined by its mean m and standard deviation s . Figure 2 shows the distributions of the portfolio LGD in the two models given that $m = 0.4$, $s = 0.15$, and $r = 0.15$. In the Tasche model we used the probability of default $p_H = 0.01$. It is obvious that the variance of the portfolio LGD is much lower in the case of Tasche model than in the case of Frye model. In fact the standard deviation of the former is approximately 4.5% while the standard deviation of the latter is 9.7%.

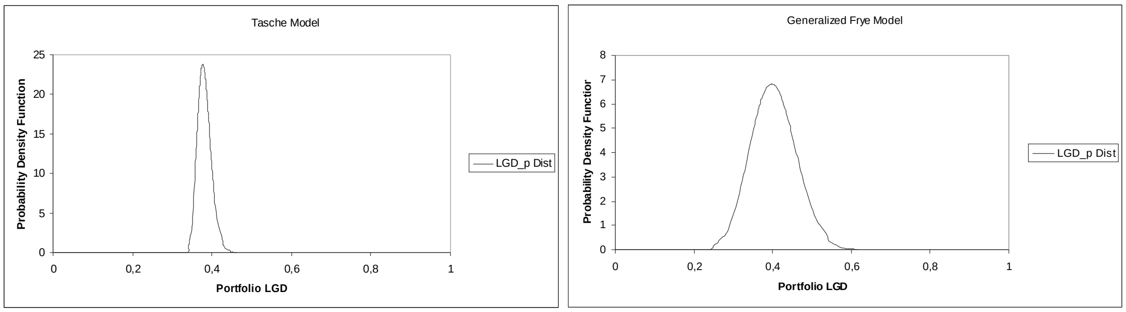


Figure 2. Portfolio loss distribution in the Tasche and generalized Frye model ($m = 0.4$, $s = 0.15$, $r = 0.15$, $p_H = 0.01$)

The Tasche model in spite of its appealing simplicity turns out to be inappropriate if unexpected loss is to be factorized according to (2). If the correlation is calibrated for the unexpected default rate calculation (i.e. around the regulatory values) then the portfolio LGD variance is too low compared to empirical observation. This follows for example from the study of Frye (2003) showing that LGD in bad years is almost twice the LGD in good years, or Frye (2000b) where the Frye model correlation coefficients r_1 and r_2 calibrated to a Moody's database appear to be almost equal. Another disadvantage of the Tasche model is that PD estimations cannot be separated from LGD estimations. On the other hand the Frye model can be calibrated separately according to volatility of frequencies of default over a number years and according to volatility of portfolio LGD observed in a time series.

4. An Analysis of the Sensitivity of the Regulatory Capital Formula

The phenomenon described in Section 2 has been partially explained in Witzany (2008) using a Merton model based simulation where we argued that a softer definition of default terminates the asset value stochastic process sooner than a hard definition of default, thus reducing the variance of losses determined by the average LGD set at the time of default.

To provide a better analytical explanation of the difference between the real loss quantile and the regulatory loss quantile estimation (and its dependence on the definition of default) we will use the Frye and Tasche one factor models as benchmarks against which we compare the regulatory unexpected loss estimation. In both the unexpected loss we need to estimate is given by

(7) $$UL = E[L \mid X = x]$$

where $x = \Phi^{-1}(a)$ and a is the regulatory probability level 0.999.

Let us consider the Tasche model first. Let $p_H = P[L > 0] = P[D_H]$ be the probability of "hard default" ($D_H \Leftrightarrow L > 0$). Note that $F_L(0) = 1 - p_H$ where F_L is the distribution function of L. Consequently $L = L(Y) = F_L^*(\Phi(Y)) = 0$ for $Y \le \Phi^{-1}(1 - p_H)$ and $L = F_L^*(\Phi(Y)) > 0$ for $Y > \Phi^{-1}(1 - p_H)$. Hence $y_H = \Phi^{-1}(1 - p_H)$ is the hard default critical point for the factor Y. As already explained banks naturally use a softer definition of default. Let us assume that such a definition of default is represented by another critical point $y < y_H$. With this new definition of default $D \Leftrightarrow Y > y$ the loss (given default) may be zero with a positive probability, $P[L = 0 \mid D] = P[y < Y \le y_H] > 0$. This new definition of default D does not change anything on the unexpected loss $UL = E[L \mid X = x]$ where the notion of default is irrelevant. On the other hand the regulatory estimation of unexpected loss turns out to be different for the two definitions of default:

(8) $$\begin{aligned} UL_{reg}^{H} &= P[Y > y_H \mid X = x] \cdot E[L \mid Y > y_H] \\ UL_{reg}(y) &= P[Y > y \mid X = x] \cdot E[L \mid Y > y]. \end{aligned}$$

Let $p = 1 - \Phi(y)$ be the probability the soft default. Note that,

(9) $$P[Y > y \mid X = x] = P[\sqrt{r}\,x + \sqrt{1-r}\,z \ > y] = \Phi\left(\frac{\Phi^{-1}(p) + \sqrt{r}\,x}{\sqrt{1-r}}\right)$$

coincides with the regulatory formula for the unexpected default rate. The difference between (8) and (7) lies in the second part of the formula (8), i.e. in $E[L \mid Y > y]$ where the regulation in general requires an average loss given default in the sense of the discussion above while the "real" unexpected loss (7) can be decomposed as

$$E[L \mid X = x] = P[Y > y \mid X = x] \cdot E[L \mid X = x, Y > y].$$

It appears obvious that $E[L \mid X = x, Y > y] > E[L \mid Y > y]$, the full proof is unfortunately rather technical. The right hand side of the inequality can be written as

$$E[L \mid Y > y] = \frac{1}{1-\Phi(y)} \int_y^\infty L(z) f(z) dz = \int_y^\infty L(z) f_1(z) dz$$

where $f(z) = \frac{1}{\sqrt{2p}} e^{\frac{-z^2}{2}}$ is the standard normal distribution density function and

$f_1(z) = \frac{f(z)}{1-\Phi(y)}$. The left hand side equals to

(10)

$$E[L \mid Y > y, X = x] = \frac{1}{1-\Phi\left(\frac{y-\sqrt{r}\,x}{\sqrt{1-r}}\right)} \int_{\frac{y-\sqrt{r}\,x}{\sqrt{1-r}}}^\infty L(\sqrt{r}\,x + \sqrt{1-r}\,z) f(z) dz =$$

$$= \frac{1}{1-\Phi\left(\frac{y-\sqrt{r}\,x}{\sqrt{1-r}}\right)} \int_y^\infty L(z) f\left(\frac{z-\sqrt{r}\,x}{\sqrt{1-r}}\right) \frac{dz}{\sqrt{1-r}} = \int_y^\infty L(z) f_2(z) dz$$

where $f_2(z) = \frac{f\left(\frac{z-\sqrt{r}\,x}{\sqrt{1-r}}\right)}{\sqrt{1-r}\cdot\left(1-\Phi\left(\frac{y-\sqrt{r}\,x}{\sqrt{1-r}}\right)\right)}$.

Both densities $f_1(z)$ and $f_2(z)$ are normalized over the interval $[y, +\infty)$ hence to show that

(11) $$\int_y^\infty L(z) f_1(z) dz < \int_y^\infty L(z) f_2(z) dz$$

we need to analyze the relationship between the two densities. It follows from properties of the normal distribution density that (provided $x > 0$ and $r > 0$) there is an $\tilde{y} > y$ so that $f_1(z) > f_2(z)$ on $[y, \tilde{y})$ and $f_1(z) < f_2(z)$ on $(\tilde{y}, +\infty)$. See Figure 3 for an illustration with $y = \Phi^{-1}(0.99)$, $x = \Phi^{-1}(0.999)$, and $r = 0.1$. Provided $L(z)$ is an increasing function (and not constant on $[y, +\infty)$) the inequality (11) follows immediately.

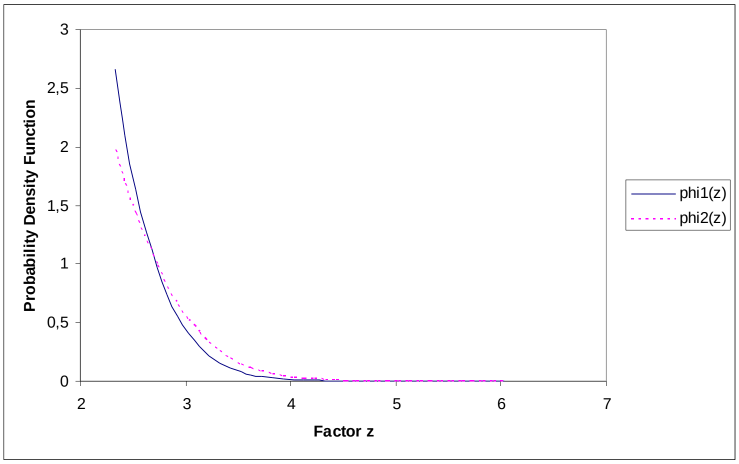


Figure 3. Density functions f_1 and f_2

To show that the regulatory unexpected loss is less than the Frye model unexpected loss is in fact much easier. In this case we just need to prove that

(12) $E[G(Y_2)] < E[G(\sqrt{r_2}X + \sqrt{1-r_2}z) \mid X = x]$

with the notation from Section 3. The left hand side simply equals to $\int_{-\infty}^{\infty} G(y)f(y)dy$ while the right hand side can be after a substitution written as $\int_{-\infty}^{\infty} G(y)f_1(y)dy$ where $f_1(y) = f\left(\frac{y-\sqrt{r}x}{\sqrt{1-r}}\right) / \sqrt{1-r}$. For $r > 0$ it can be verified that the function $f_1(y) < f(y)$ on an interval $(-\infty, y_0)$ and $f_1(y) > f(y)$ on $(y_0, +\infty)$, see Figure 4. Consequently again (12) holds provided G is non-decreasing and strictly increasing on a non-trivial interval.

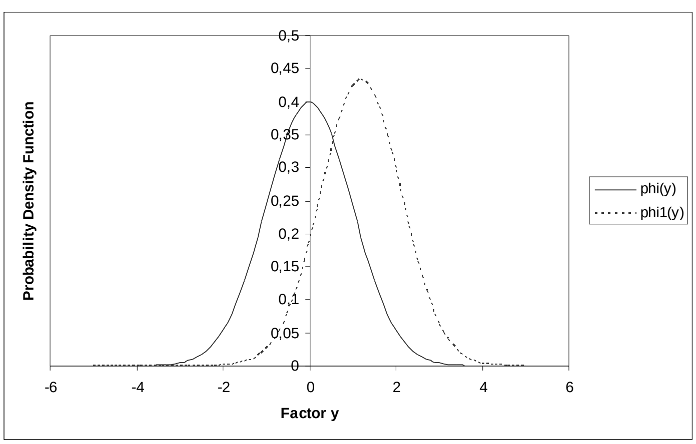


Figure 4. Density functions f and f_1

Next we want to show that the function $UL_{reg}(y) = P[Y > y \mid X = x] \cdot E[L \mid Y > y]$ defined according to (8) is an increasing function of $y \le y_H$ for certain range of feasible values for y and r. Since $E[L \mid Y > y] = \frac{E[L]}{P[Y > y]}$ we just need to show that the ratio between the unexpected loss $UL_{reg}(y)$ and the expected loss $E[L]$ not depending on y

$$(13) \qquad h(y, r) = \frac{UL_{reg}(y)}{E[L]} = \frac{P[Y > y \mid X = x]}{P[Y > y]} = \frac{1 - \Phi\left(\frac{y - \sqrt{r}\,x}{\sqrt{1-r}}\right)}{1 - \Phi(y)}$$

is an increasing function of y. Note that the equation (13) is identical for the Tasche and Frye model. Unfortunately we cannot prove generally that the function $h(y, r)$ is increasing in $y > 0$ for any given correlation $r > 0$. In fact it is not increasing on $(0, \infty)$ with $r > 0$ since $\frac{y - \sqrt{r}\,x}{\sqrt{1-r}} > y$ and so $h(y, r) < 1$ for large y while clearly $h(y, r) > 1$ for smaller values of $y > 0$. However it can be shown using numerical approximations that the function is increasing over a range of admissible values for y and r. Figure 5 shows the function (13) strongly increasing with the values $x = \Phi^{-1}(0.999)$ and $r = 0.4$, $r = 0.1$, $r = 0.15$, and for $y = \Phi^{-1}(1 - p_D)$ over the range [0.8,3.7] corresponding to admissible PD values in the interval [0.01%,21.2%].

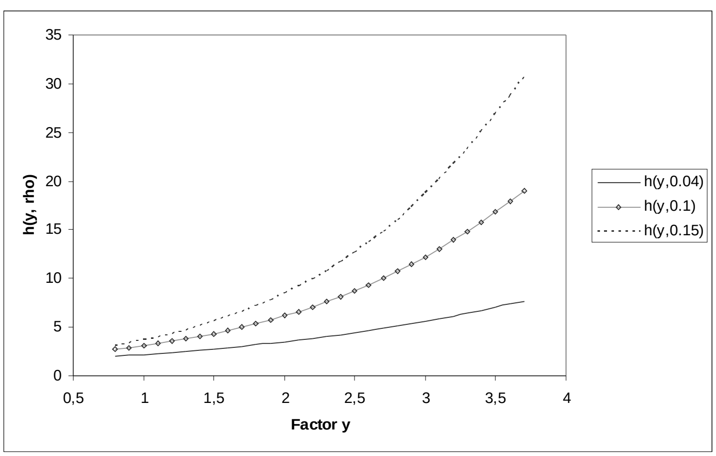


Figure 5 The ratio between the unexpected loss $UL_{reg}(y)$ and the expected loss $E[L]$ increases with y

5. Improved Regulatory Formula

In Section 3 we gave a general definition of one (systemic) factor model. We have seen that if D is a consistent notion of default then the loss may be decomposed to $UL = P[D \mid X = x] \cdot E[L \mid D, X = x]$. It is not in general obvious that the conditional default rate $P[D \mid X = x]$ as well as the conditional loss given default $E[L \mid D, X = x]$ are increasing functions of x. However it is a property of the aforementioned one-factor models (Tasche, Frye, Pykhtin). Consequently it is correct in the context of one-factor models where both the conditional PD and conditional LGD are increasing functions of the systemic factor X to state that

(14) $$UL = UDR \cdot ULGD$$

where $UDR = P[D \mid X = x]$ is the a - quantile of possible default rates and $ULGD = E[L \mid D, X = x]$ the a - quantile of possible LGDs with x being the a - quantile of X. The unexpected default rate is estimated consistently by the reasonably well by the regulatory formula (1). But we improve it significantly requiring that LGD is not the expected loss given default but the unexpected portfolio level loss given default (ULGD) on the 99.9% probability level.

For practical applications we propose to use instead of the generalized Frye's model. In the notation of Section 3 we just need to evaluate

(15) $$ULGD = E[L \mid Y_1 > y_D, X = x] = E[G(Y_2) \mid X = x] = \int_{-\infty}^{+\infty} G(\sqrt{r_2}x + \sqrt{1 - r_2}z)f(z)dz$$

To complete our model we need to propose an appropriate loss given default function G. We may follow Witzany (2008) specifying that $G(Y)$ has a beta distribution calibrated to empirical mean and standard deviation. However since we consider that the default definition may be in a practice a softer one with a non-negligible percentage p_{cure} of receivables marked as defaulted being cured, i.e. ultimately ends up with zero loss, we extent the model as follows. Let m and s be the mean and standard deviation of observed positive losses assumed to have a beta distribution with minimum 0 and maximum 1. Let $B(t, m, s)$ be the corresponding cumulative Beta distribution function on [0,1] then the mixed distribution function incorporating the possibility of cures is defined by

(16) $\quad F(t) = p_{cure} + (1 - p_{cure}) \cdot B(t, m, s)$.

Finally setting $G(Y) = F^*(\Phi(Y))$ we see that G has the distribution given by F . To estimate the 99.9% ULGD we just need to evaluate (15) numerically for $x = \Phi^{-1}(0.999)$. Figure 6 illustrates the account-level LGD density function given by F for a given set of parameters (with mass weight p_{cure} at 0) and the transformed portfolio level LGD density function of $E[L \mid X]$ derived from (15).

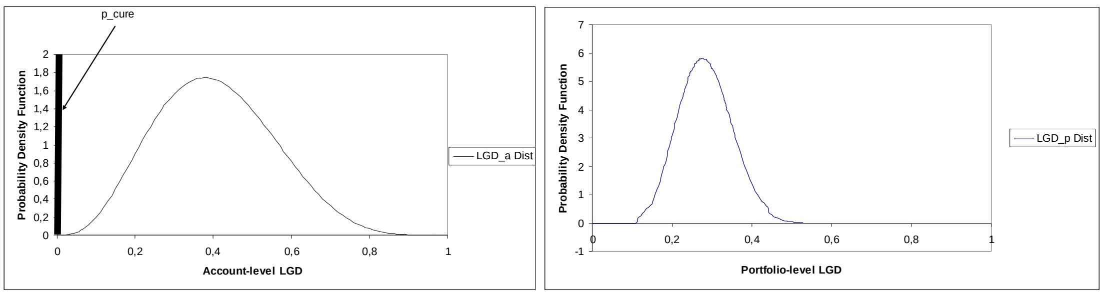


Figure 6. Account and transformed portfolio loss given default distribution ($p_{cure} = 0.3$, $m = 0.4$, $s = 0.15$, $r = 0.1$).

Unexpected loss estimated using the described technique is however still sensitive to the definition of default although Figure 8 shows that the sensitivity is moderate and opposite compared to the regulatory capital (unexpected loss estimation increases with softer definition of default). Another applicable solution is then to adjust the probability of (conventional soft) default p with the observed probability of cures p_{cure} and then apply the hard default based formula

(17) $\quad UL = UDR(p(1 - p_{cure})) \cdot ULGD(m, s)$

where unexpected loss given (hard default) is estimated according to (15) using the beta distribution with mean m and standard deviation s .

6. Numerical Examples

We are going to compare the values of regulatory unexpected loss in different scenarios: unexpected loss in the Tasche model, and unexpected loss in the Frye model with and without the cure rate. The scenarios are specified by the probability of hard default p_H , loss given (hard) default mean m and standard deviation s , correlations $r = r_1$ and r_2 , and the cure rate p_{cure}. The probability of soft default is then recalculated as $p_S = p_H / (1 - p_{cure})$, the loss given soft default mean as $m_S = m(1 - p_{cure})$ and the standard deviation

$$s_S = \sqrt{(1 - p_{cure})(s^2 + p_{cure} m^2)}.$$

Figure 7 shows how it is difficult to align the Tasche and Frye model. If we fix $r = 15\%$ as the correlation related to unexpected default rate and the other parameters as specified below than in order to obtain the same unexpected loss in the Frye model as in the Tasche model the LGD correlation must be reduced down to 1% or even less. Such a calibration is n contradiction with empirical studies like Frye (2000b, 2003). Thus we focus rater on the Frye modeling approach.

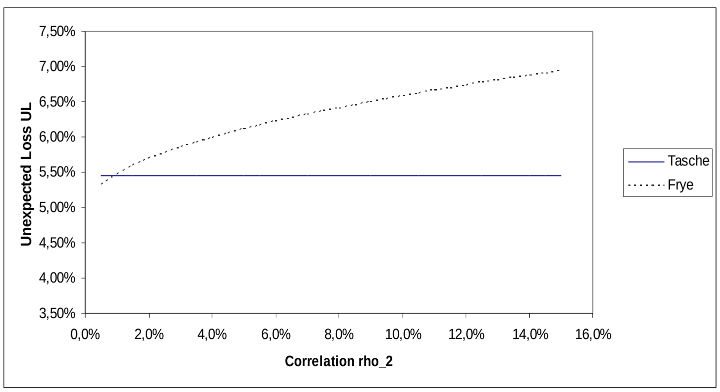


Figure 7. Comparison of the 99,9% unexpected loss in the Tasche model and Frye model depending on r_2 (with $p_H = 1\%$, $m = 45\%$, $s = 15\%$, $r = 15\%$, $p_{cure} = 0\%$)

Figure 8 on compares the regulatory unexpected loss estimate with different estimation approaches based on the Frye model explained at the end of the previous section. While the UL_reg curve shows the regulatory unexpected loss declining with the cure rate going up, the curve UL_Frye_S based on the beta distribution calibrated to m_S and s_S turns out to be increasing. The dependence is weaker if we use the mixed beta distribution (16) (UL_Frye_S_c) and logically it is fully eliminated (UL_Frye_H) when we use (17). We consider the positive sensitivity of UL_Frye_S_c on the cure rate motivating banks to use a harder definition of default to be much more acceptable than the negative sensitivity of UL_reg which motivates banks to use a softer definition of default that is not usually ideal for credit risk modelling as pointed out in Witzany (2008). The problem is fully solved by recalculating the probability of soft default to the probability of hard default, which might be however a little bit difficult to communicate in practice.

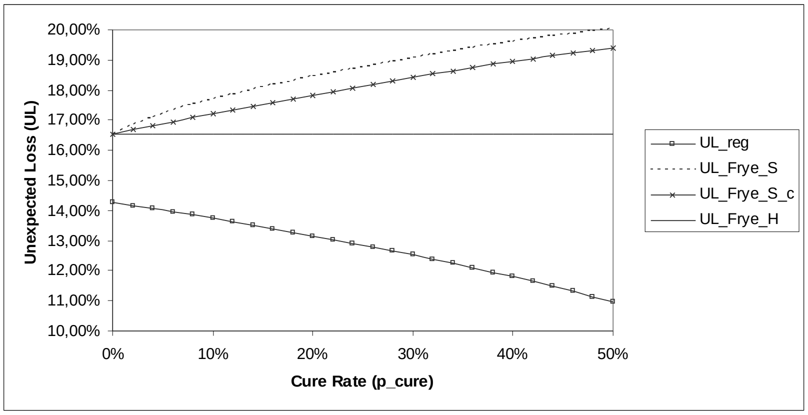
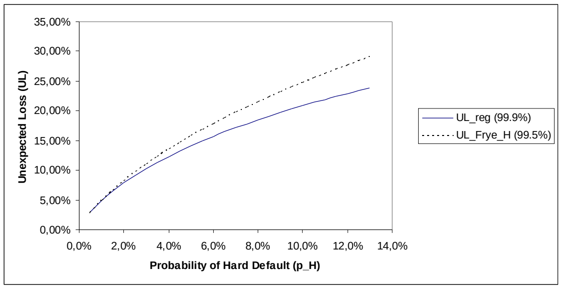


Figure 8. Regulatory and Frye model based 99,9% unexpected loss estimations depending on p_{cure} (with $p_H = 2.5\%$, $m = 70\%$, $s = 15\%$, $r = r_1 = 15\%$, $r_2 = 8\%$)

Incorporation of unexpected loss given default into unexpected loss calculation significantly increases the value compared to the regulatory unexpected loss. If we wanted to setup the model in line with current regulatory capital values we could consider reducing of the (artificially high) regulatory level. It turns out that the level of 99.5% (proposed e.g. for the Solvency II) leads to comparable values of the regulatory UL on the 99.9% level and the Frye model UL on the 99.5% level. The relationship nevertheless depends on the s and r_2 values. Figure 9 finally compares the sensitivity the 99.9% regulatory UL and 99.5% Frye model UL (17) on the probability of default and expected loss given default, other parameters fixed. The 99.5% Frye UL turns out to be more sensitive than the regulatory UL with respect to the probability of default but less to the expected loss given default. Hence by an appropriate recalibration of the confidence level we do not obtain the same unexpected loss estimations in all scenarios but using the proposed model we obtain a better correspondence between the real risk and the economic capital, more robust calculations, and at the same time overall same average level of capital.



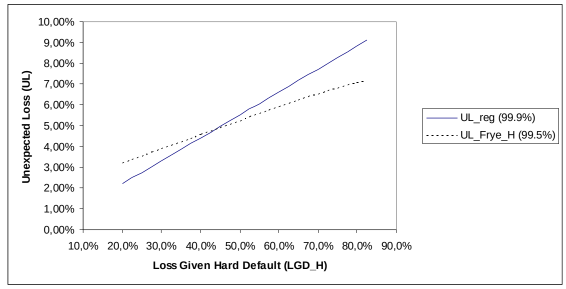


Figure 9. Sensitivity of the 99.9% regulatory and 99.5% Frye model unexpected loss on p_H and m (with, $s = 22\%$, $r = r_1 = 15\%$, $r_2 = 13\%$, $m = 45\%$ in the first graph, and $p_H = 1\%$ in the second)

7. Conclusion

We have demonstrated and analytically explained that the regulatory capital according to the Basle II formula is sensitive to the definition of default. We have shown that the problem may be relatively simply solved in the context of general single factor models requiring that the LGD parameter is reinterpreted as the 99.9% percentile of possible losses given default. We have considered three particular one (systemic) factor models and concluded that the one with two idiosyncratic factors proposed by Frye is the most appropriate to implement practically our model. The best results are provided by the model where the observed probability of soft default is adjusted using the cure rate to obtain the probability of hard default (which can be fully determined only ex post). Since the extended model gives higher unexpected loss values the confidence level can be recalibrated to a lower value (e.g. 99.5%) to achieve comparable capital levels. The resulting formula, which could replace the regulatory formula, provides a more robust and economically more faithful estimates of unexpected credit losses.

Literature

Acharya, Viral, V., S. Bharath and A. Srinivasan, 2007, "Does Industry-wide Distress Affect Defaulted Firms? – Evidence from Creditor Recoveries," Journal of Financial Economics 85(3):787–821.

Altman E., Resti A., Sironi A., 2004, "Default Recovery Rates in Credit Risk Modelling: A Review of the Literature and Empirical Evidence," Economic Notes by Banca dei Paschi di Siena SpA, vol.33, no. 2-2004, pp. 183-208

E. Altman – G. Fanjul, 2004, ''Defaults and Returns in the High Yield Bond Market: Analysis through 2003'', NYU Salomon Center Working Paper

Basle Committee on Banking Supervision Paper, 2005, "Guidance on Paragraph 468 of the Framework Document"

Basel Committee on Banking Supervision, 2006, **"**International Convergence of Capital Measurement and Capital Standards, A Revised Framework – Comprehensive Version", Bank for International Settlements, 2006

Brady B., Chang P., Miu P., Ozdemir B., Schwartz D., 2006, "Discount Rate for Workout Recoveries: An Empirical Study", working paper

Dullman, K. and M. Trapp, 2004, "Systematic Risk in Recovery Rates – An Empirical Analysis of U.S. Corporate Credit Exposures", EFWA Basel Paper.

K. Emery, 2003, Moody's Loan Default Database as of November 2003, Moody's Investors Service.

Frye, J., 2000a, "Collateral Damage", RISK 13(4), 91–94.

Frye, J., 2000b, "Depressing recoveries", RISK 13(11), 106–111.

Frye, J., 2003, “A false sense of security, RISK 16(8), 63–67.

Gordy, M. B., 2003, “A risk-factor model foundation for ratings-based bank capital rule”, Journal of Financial Intermediation, Vol 12, pp. 199-232.

G. Gupton, D. Gates, and L. Carty, 2000, “Bank loan losses given default”, Moody’s Global Credit Research, Special Comment.

Gupton G.M., 2005, „Advancing Loss Given Default Prediction Models: How the Quiet Have Quickened”, Economic Notes by Banca dei Paschi di Siena SpA, vol.34, no. 2-2005, pp. 185-230

Kim J., Kim H., 2006, „Loss Given Default Modelling under the Asymptotic Single Risk Factor Assumption”, Working Paper, MPRA

Merton, Robert C., 1974, “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” The Journal of Finance, Vol. 29, pp. 449-470.

Pykhtin, M., 2003, “Unexpected recovery risk”, Risk, Vol 16, No 8. pp. 74-78.

Schuermann T., 2004, “What Do We Know About Loss Given Default”, Credit Risk Models and Management, 2nd Edition, London, Risk Books

Tasche, Dirk., 2004, “The single risk factor approach to capital charges in case of correlated loss given default rates”, Working paper, Deutsche Bundesbank, February 2004.

Vasicek O., 1987, “Probability of Loss on a Loan Portfolio,” KMV Working Paper

Witzany, J., 2008, “Basle II Capital Requirement Sensitivity to the Definition of Default”, Working paper, University of Economics (Prague)

IES Working Paper Series

2009

1. František Turnovec : *Fairness and Squareness: Fair Decision Making Rules in the EU Council?*
2. Radovan Chalupka : *Improving Risk Adjustment in the Czech Republic*
3. Jan Průša : *The Most Efficient Czech SME Sectors: An Application of Robust Data Envelopment Analysis*
4. Kamila Fialová, Martina Mysíková : *Labor Market Participation: The Impact of Social Benefits in the Czech Republic*
5. Kateřina Pavloková : *Time to death and health expenditure of the Czech health care system*
6. Kamila Fialová, Martina Mysíková : *Minimum Wage: Labour Market Consequences in the Czech Republic*
7. Tomáš Havránek : *Subsidy Competition for FDI: Fierce or Weak?*
8. Monika Slabá : *Liberalisation of Natural Gas Market – EU Vision vs. Reality*
9. Jiří Witzany : *Loss, Default, and Loss Given Default Modeling*

All papers can be downloaded at: http://ies.fsv.cuni.cz

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz