Abstract

We adapt structural models of default risk to take into account the special nature of bank assets. The usual assumption of log-normally distributed asset values is not appropriate for banks. Bank assets are risky debt claims, which implies that they embed a short put option on the borrowers’ assets, leading to a concave payoff. This has important consequences for banks’ risk dynamics and distance to default estimation. Due to the payoff non-linearity, bank asset volatility rises following negative shocks to asset values. As a result, standard structural models in which the asset volatility is assumed to be constant can severely understate banks’ default risk in good times when asset values are high. Bank equity payoffs resemble a mezzanine claim rather than a call option. Bank equity return volatility is therefore much more sensitive to negative shocks to asset values than in standard structural models.
I Introduction

The distress that many banks experienced during the recent financial crisis has brought renewed emphasis on the importance of understanding and modeling bank default risk. Assessment of bank default risk is important not only for investors in banks’ debt securities, but also for risk managers analyzing counterparty risks and for regulators gauging the risk of bank failure. Accurate modeling of bank default risk is also required for valuing the benefits that banks derive from implicit and explicit government guarantees.

In many applications of this kind, researchers and analysts rely on structural models of default risk in which equity and debt are viewed as contingent claims on the assets of the firm. Following Merton (1974), the standard approach (which we call the Merton model) is to assume that the value of the assets of the firm follows a log-normal process, which allows to use price the options embedded in the firm’s equity and debt as in Black and Scholes (1973). In some cases, the distance to default computed from this model is used as one of the ingredients to empirically model bank default risk. Recent examples of bank default risk analysis based on the Merton model includes Acharya, Anginer, and Warburton (2014) and Schweikhard, Tsesmelidakis, and Merton (2014), who study the value of implicit (too-big-to-fail) government guarantees. There is also an extensive literature that has applied structural models to price deposit insurance, going back to Merton (1977), Marcus and Shaked (1984), and Pennacchi (1987).

The Merton model’s assumption of log-normally distributed asset values may provide a useful approximation for the asset value process of a typical non-financial firm. However, for banks this assumption is clearly problematic. Much of the assets of a bank consist of debt claims such as mortgages that involve contingent claims on the assets of the borrower. The fact that the upside of the payoffs of these debt claims is limited is not consistent with the unlimited upside implied by a log-normal distribution.

In this paper, we propose a modification of the Merton model that takes into account the debt-like payoffs of bank assets. Our approach to deal with this problem is to apply the log-normal distribution assumption not to the assets of the bank, but to the assets of the bank’s borrowers. More precisely, we model banks’ assets as a pool of loans where borrowers’ loan repayments depend on the value of their assets at loan maturity as in Vasicek (1991). Borrower asset values are subject
This options-on-options feature of bank equity and debt has important consequences for the implied default risk and equity risk dynamics. To illustrate the main intuition, it is useful to consider the simplified case without idiosyncratic risk in which all borrowers are identical with perfectly correlated defaults. Assume further that the maturity of the (zero-coupon) debt issued by the bank equals the maturity of the (zero-coupon) loans made by the bank. In this case, the payoffs at maturity as a function of borrower asset value are as shown in Figure 1. In this example, the borrowers have loans with face value 0.80 and the bank has issued debt with face value 0.60. Since the maximum payoff the bank can receive from the loans is their face value, the bank asset value is capped at 0.80. Only when borrower assets fall below 0.8 is the bank asset value sensitive to borrower asset values.

Clearly, the bank asset value cannot follow a log-normal distribution (which would imply unlimited upside). Since the bank’s borrowers keep the upside of a rise in their asset value above the loan face value, the bank’s equity payoff does not resemble a call option on an asset with unlimited...
upside, but rather a mezzanine claim with two kinks. This nature of the banks’ asset payoffs is in conflict with the assumptions of the standard Merton model.

This mezzanine-like nature of the bank’s equity claim has important consequence for the risk dynamics of bank equity and for the distance-to-default estimation. Due to the capped upside, bank volatility will be very low in “good times” when asset values are high and it is likely that asset values at maturity will end up towards the right side in Figure 1 where the bank’s equity payoff is insensitive to fluctuations in borrower asset values.

A standard Merton-model in which equity is a call option on an asset with constant volatility misses these nonlinear risk dynamics. Viewed through the lens of this standard model it might seem that a bank in times of high asset values is many standard deviations away from default. But this conclusion would be misleading because it ignores the fact that volatilities could rise dramatically if asset values fall. Similarly, the standard model would give misleading predictions about the riskiness of bank assets, equity, and debt.

Going beyond this simple illustrative example, our model incorporates additional features such as idiosyncratic borrower risks, staggered remaining loan maturities, and the replacement of maturing loans with new loans. While these generalizations are important to get a more realistic model that we can use for a quantitative evaluation of bank risk, the addition of these features does not change the basic insight about the mezzanine-claim nature of bank equity and the resulting consequences for risk dynamics and distance to default.

The assess the differences between our modified model and the Merton model, we simulate data from our modified model and ask to what extent an analyst using the Merton model would mis-judge the risk-neutral probability of default. This error is particularly stark when asset values are high relative to the face value of the bank’s debt. In this case, bank asset payoffs are likely to stay in the flat region in Figure 1 and bank equity payoffs are also likely in the flat region. As a consequence, equity volatility is low. Based on the Merton model, the analyst would infer from low equity volatility that asset volatility must be low. Furthermore, since asset volatility is constant in the Merton model, the analyst would (wrongly) assume that asset volatility will remain low at this level in the future. What the Merton model misses in this case is that asset volatility could rise substantially following a bad asset value shock, because the bad shock brings the likely asset payoff closer to or into the downward sloping region in Figure 1. As a result, the Merton
model substantially overestimates the distance to default and it underestimates the risk-neutral probability of default.

We then calibrate our modified model and the standard Merton model to quarterly bank panel data from 2002 to 2012. We choose the value and volatility of bank assets (in the case of the Merton model) or borrower assets (in case of our modified model) to match the observed market value of equity and its volatility. Even though both models are calibrated to the same data, their implied risk-neutral default probabilities are strikingly different. In line with the simulations we discussed above, the differences are particularly big in the years before the financial crisis when equity values were high and volatility low. Based on the Merton model, the risk-neutral default probability of the average bank in 2006 over a 5-year horizon is less than 10%. In contrast, the risk-neutral default probability from our modified model is two to three times as high. Translated into credit spreads, this would imply an annualized spread of around 10 basis points in the Merton model and close to 100 basis points according to our modified model.

Once the financial crisis hit in 2007/08, the models’ predictions are not so different anymore. At this point, bad asset value shocks had moved banks into the downward-sloping asset payoff region in Figure 1. In this region, the kink in the asset payoff becomes less relevant and the predictions from our modified model are close to those from the standard Merton model. In periods of the most extreme distress, Merton model default probabilities can even exceed those from our modified model.

Thus, the key problem with applications of standard structural models to banks is that they understate the risk of default in “good times.” This is an important issue, for example, for the estimation of the value of explicit or implicit government guarantees. Based on a standard Merton model calibrated to data from pre-crisis times, one may be lead to the conclusion that the value of a guarantee is almost nil when, in fact, the value is a lot if one takes into account that banks’ asset volatility will go up when asset values fall.

Our paper relates to several strands in the literature. Gornall and Strebulaev (2014) model bank assets as loan portfolios in similar ways as we do, albeit without staggering of loan maturities. Their focus is on modeling bank’s capital structure choices in equilibrium, while we focus on implications for default risk estimation and valuation of bank’s securities. Duffie, Jarrow, Purnanandam, and Yang (2003) develop a reduced-form pricing approach for pricing of deposit insurance. The reduced-
form approach permits a lot of flexibility to obtain realistic default risk estimates, but the structural approach that we pursue here is useful for understanding the economic drivers of default risk (which may in turn be useful to develop well-specified reduced-form models.) Kelly, Lustig, and Van Nieuwerburgh (2011) estimate the value of implicit government guarantees for the banking system by comparing prices of options on a banking index and a portfolio of options on individual bank stocks. Their method involves fitting models with stochastic volatility and jumps to option prices. In these models, the correlation between returns and shocks to volatility (the “leverage” effect) is a reduced-form parameter. Our structural model of bank risk predicts a specific (non-linear) relation between bank equity returns and bank equity volatility.

The rest of the paper is organized as follows. Section II presents our modified model and simulations to illustrate the key differences to the standard Merton model. In Section III we apply the model to empirical bank panel data and we analyze the resulting estimates of default risk. Section IV concludes.

II Structural Model of Default Risk for Banks

Unlike the simplified case in Figure 1, we now set up a model in which borrower asset have idiosyncratic risk and banks issue loans with staggered maturities. Both of these additional features are important because they lead to some smoothing of the bank asset payoff function in Figure 1.

Consider a setting with continuous time. A bank issues zero-coupon loans with maturity $T$. Loans are issued in staggered fashion to $N$ cohorts of borrowers. Cohorts are indexed by the remaining maturity on their loans at $t = 0$, $\tau = T, T(N - 1)/N, \ldots, 1/N$. These loans were issued at $\tau - T$ with face value $F_1$. Each cohort is comprised of a continuum of borrowers indexed by $i \in [0, 1]$ with mass $1/N$.

Let $A_{\tau,i}^t$ denote the collateral value of a borrower $i$ in cohort $\tau$ at time $t$. Under the risk-neutral measure, the asset value evolves according to the stochastic differential equation

$$
\frac{dA_{\tau,i}^t}{A_{\tau,i}^t} = (r - \delta)dt + \sigma(\sqrt{\rho}dW_t + \sqrt{1 - \rho}dZ_{\tau,i}^t),
$$

where $W$ and $Z_{\tau,i}^t$ are independent standard Brownian motions, $\delta$ is a depreciation rate, and $r$ is
risk-free rate. The $Z^{τ,i}$ processes are idiosyncratic and independent across borrowers. This is a one-factor model of borrower asset values as in Vasicek (1991). The parameter $\rho$ represents the correlation of asset values that arises from common exposure to $W$.

At the time of the initial loan issue, $τ - T$, borrowers’ collateral is worth $A^{τ,i}_{τ} = 1$. Thus, the initial loan-to-value ratio is

$$\ell = F_1 e^{-\mu T},$$  

with $\mu$ as the promised yield on the loan (that we will solve for below). In line with standard structural models of credit risk, we assume that borrowers default if the asset value at maturity is lower than the amount owed. The payoff at maturity $t = τ$ to the bank from an individual borrower’s loan then is

$$L^{τ,i}_{τ} = \min(A^{τ,i}_{τ}, F_1).$$  

To analyze the payoff to the bank from the whole loan portfolio, it is useful to first solve for the aggregate value of collateral in cohort $n$, which is,

$$A^{τ}_{τ} = \frac{1}{N} \int_{0}^{1} A^{τ,j}_{τ} dj$$

$$= \frac{1}{N} \exp \left\{ (r - \delta)T - \frac{1}{2} \rho \sigma^2 T + \sigma \sqrt{\rho} (W_τ - W_{τ-T}) \right\}, \quad (4)$$

and the aggregate log asset value, which is

$$a^{τ}_{τ} = \frac{1}{N} \int_{0}^{1} \log A^{τ,j}_{τ} dj$$

$$= \frac{1}{N} (r - \delta)T - \frac{1}{2} \sigma^2 T + \sigma \sqrt{\rho} (W_τ - W_{τ-T}). \quad (5)$$

Since idiosyncratic risk fully diversifies away with a continuum of borrowers in each cohort, the stochastic component of the aggregate asset value a cohort depends only on the common factor realization $W_τ - W_{τ-T}$.

We now obtain the payoff that the bank receives at maturity from the portfolio of loans given
to cohort \( \tau \) as

\[
L^*_\tau = \frac{1}{N} \int_0^1 L^*_{\tau,j} \,dj \\
= \frac{1}{N} \int_0^1 A^*_{\tau,j} \,dj - \frac{1}{N} \int_0^1 \max(A^*_{\tau,j} - F_1, 0) \,dj \\
= \frac{1}{N} \left[ A^*_\tau \Phi(d_1) + F_1 \Phi(d_2) \right]
\]

where the last equality follows from the properties of the truncated log-normal distribution, \( \Phi(.) \) denotes the standard normal CDF, and

\[
d_1 = \frac{\log F_1 - a^*_\tau}{\sqrt{1 - \rho} \sqrt{T} \sigma} - \sqrt{1 - \rho} \sqrt{T} \sigma, \\
d_2 = -\frac{\log F_1 - a^*_\tau}{\sqrt{1 - \rho} \sqrt{T} \sigma}.
\]

The \( \max(A^*_{\tau,j} - F_1, 0) \) term in (6) reflects the option value for the borrower, i.e., the upside of the collateral value that is retained by the borrower. Conditional on \( W_\tau - W_{\tau-T} \), there are some borrowers in cohort \( \tau \) for whom this option is in the money, and others for whom it is not, depending on the realization of the idiosyncratic shock. This is why \( d_1 \) and \( d_2 \) are functions of idiosyncratic risk \( \sqrt{1 - \rho} \sqrt{T} \sigma \). Thus, while idiosyncratic risk is diversified away in the aggregate borrower asset value, it matters for loan payoffs, because borrower default depends on idiosyncratic risk.

We assume that loans are priced competitively, and so the promised yield on the loan can now be found as the \( \mu \) that solves

\[
F_1 e^{-\mu T} = e^{-r T} \mathbb{E}^Q_{\tau-T}[L^*_\tau].
\]

where \( \mathbb{E}^Q[.] \) denotes a conditional expectation under the risk-neutral measure. The risk-neutral expectation on the right-hand side can be evaluated by simulating the distribution of \( L^*_\tau \), based on (6) and the simulated distribution of \( A^*_n \), under the risk-neutral measure.

At \( t = \tau \), the bank fully reinvest the proceeds from the maturing loan portfolio of cohort \( \tau \), \( L^*_\tau \), into new loans, with uniform amounts, to members of the same cohort. The new loans carry a face value of

\[
F_2 = L^*_\tau e^{\mu T}.
\]
We assume that the bank keeps the time-of-issue loan-to-value ratio at the same level, i.e., \( \ell \), as for the initial round of loans. Borrowers reduce or replenish collateral assets accordingly: the asset value of each member of cohort \( \tau \) is reset to the same value

\[
A_{\tau,i}^{\tau,i} = \frac{L_{\tau}^{\tau,i}}{\ell}.
\]  

(11)

an instant after the re-issue of the loans. The cohort-level aggregates \( A_{\tau}^{\tau} \) and \( a_{\tau}^{\tau} \) for \( t > \tau \) are based on these re-initialized asset values.

The aggregate payoff of the portfolio of loans of cohort \( \tau \) at the subsequent maturity date \( \tau + T \) then follows, along similar lines as above, as

\[
L_{\tau+T}^{\tau} = \frac{1}{N} \int_{0}^{1} L_{\tau+T}^{\tau,j} dj
\]

\[
= \frac{1}{N} \int_{0}^{1} A_{\tau+T}^{\tau,j} dj - \frac{1}{N} \int_{0}^{1} \max(A_{\tau+T}^{\tau,j}, 0) dj
\]

\[
= \frac{1}{N} \left[ A_{\tau+T}^{\tau} \Phi(d_3) + F_2 \Phi(d_4) \right]
\]  

(12)

where

\[
d_3 = \frac{\log F_2 - a_{\tau+T}^{n}}{\sqrt{1 - \rho} \sqrt{T \sigma}} - \frac{1}{\sqrt{1 - \rho} \sqrt{T \sigma}},
\]  

(13)

\[
d_4 = \frac{- \log F_2 - a_{\tau+T}^{n}}{\sqrt{1 - \rho} \sqrt{T \sigma}}.
\]  

(14)

Thus, after the roll-over into new loans, there are two state variables to keep track of: \( A_{\tau+T}^{\tau} \) and \( F_2 \) depend on: The change of the common factor since roll-over, \( W_t - W_{\tau} \), and \( L_{\tau}^{\tau} \) (which in turn is driven by \( W_t - W_{\tau-T} \)).

The payoffs in (12) and (6) together allow us to describe the distribution of the bank’s assets. Consider, for example, the aggregate value of the bank’s loan portfolio at \( t = H \), where \( H < T \). Aggregating across all loans outstanding at this time, we get

\[
V_H = \sum_{\tau < H} e^{-r(\tau+T-H)} E_H^Q[L_{\tau+T}^{\tau}] + \sum_{\tau \geq H} e^{-r(\tau-H)} E_H^Q[L_{\tau}^{\tau}]
\]  

(15)

where the first term aggregates over cohorts whose loans have been rolled over into a second round,
Figure 2: Bank asset value at bank debt maturity as a function of aggregate borrower asset value. Simulated bank asset values shown as dots. Dashed line shows the kinked payoff that would result with perfectly correlated borrower asset values and without staggering of loan maturities while the term aggregates over cohorts that still have the initial first-round loans outstanding. Substituting in from (6) and (12) yields an expression in which the only source of stochastic shocks is the common factor $W$. Therefore, by simulating $W$ we can simulate the distribution of $V_H$ and price contingent claims whose payoffs are functions of $V_H$.

Figure 2 shows the simulated bank asset value, $V_H$, based on 10,000 draws of the common factor paths plotted against the aggregate borrower asset value at $t = H$. Parameters are set at $H = 5$, $T = 10$, $\sigma = 0.2$, $\rho = 0.5$, $r = 0.01$, $\delta = 0.005$, and $F_1 = 0.8$. Unlike in Figure 1, there is no sharp kink in the bank’s asset value in Figure 2. There are two reasons for the lack of sharp kink. First, in this case, loans in the bank’s portfolio are not at maturity. For $H < \tau \leq T$, loans have not matured yet, while for $0 \leq \tau < H$, they have been rolled over into new loans. Second, the existence of idiosyncratic borrower risk makes the borrower’s default option more valuable and the loan less valuable to the bank, particularly when the asset value is close to the face value of the debt.

Moreover, unlike in Figure 2, there is dispersion in the bank asset value conditional on aggregate borrower asset value. The reason is that for loans that have been rolled over into a second generation
of loans, the face value of the loan depends on the common factor realization up to the roll-over date $\tau$. For example, if $W_\tau$ is low, there will be more defaults at $\tau$ and hence the face value of new loans will be lower than if $W_\tau$ is high.

$V_0$, the total market value of the bank’s assets at $t = 0$ can be computed as

$$V_0 = e^{-rH}E^Q_0[V_H]$$

Now suppose the bank has issued zero-coupon debt maturing at $t = H$ with face value $D$. Similar to standard structural models we assume that the bank will pay off its creditors in full if there are sufficient assets available to do so. The bank will default if the asset value at maturity is lower than the face value of debt. To allow for a realistic calibration in our empirical exercise below, we also introduce payouts of the bank to its shareholders. For simplicity, we introduce them as a single certain payment,

$$Y_H = F_1 \gamma H,$$

just before the debt matures. The assets that leave the bank through these payouts are no longer
available to pay off the debt holders at $t = H$. The value of the bank’s creditors’ claim at $t = 0$ then is

$$B_0 = e^{-rH} D - e^{-rH} E^Q_t[(D - V_H + Y_H)^+]$$  \hfill (18)

The bank’s equity value, without the claim to the dividends paid just before maturity, then follows as

$$S_0 = V_0 - e^{-rH} Y_H - B_0$$  \hfill (19)

Figure 4 shows simulation results for the relationship between $S_0$ and $V_0$. We use the same parameters as in Figure 2 and $D = 0.6$, $Y_H = 0.008$. As Figure 3 shows, the value of bank equity is concave in bank assets for large values. This is in contrast to the standard Merton model in which the equity value asymptotes towards a slope of one. Thus, we get a similar mezzanine-like payoff for bank equity as in our preliminary analysis in Figure 1, albeit with smoothed kinks.

Figure 4 shows the instantaneous volatility of the bank’s equity. Given knowledge of the parameters, one can numerically compute the instantaneous equity volatility from the first derivative of $S_0$ with respect to $W_0$. Since common factor shocks are the only source of stochastic shocks to
the bank asset value, this derivative with respect to \( W_0 \) is directly related to the slope of the curve in Figure 3. As Figure 4 shows, equity volatility converges towards zero for high bank asset values. In the Merton model, in contrast, equity volatility would asymptote towards a fixed asset volatility.

This very low equity volatility at high asset values arises from the mezzanine-claim nature of bank equity. Positive shocks to asset values raise borrower asset values far above the default thresholds. As a result, bank assets have very low instantaneous risk in good times and bank equity risk resembles the risk of a bond because the region of concavity dominates. Application of a standard Merton model with log-normal asset value would miss this non-linearity in bank’s equity risk. Our modified model suggests that this non-linearity is a key property of bank equity risk dynamics.

In particular, our model makes clear that low instantaneous bank equity volatility in good times can quickly turn into high risk in bad times if asset values fall. In the standard model with a log-normal asset value, the rise in bank equity volatility with a fall in asset values is moderate because bank asset volatility is fixed. In our modified model, bank asset volatility goes up as loans fall in value and become riskier. The rise in equity volatility following bad shocks is therefore more dramatic than in the standard Merton model.

II.A Default Risk Assessment: Comparison with Standard Merton Model

The highly nonlinear relation between bank asset value and bank equity risk due to the short put option embedded in bank assets leads to important consequences for distance to default estimation and empirical assessment of default risk with structural models. To illustrate these consequences, we now analyze a setting in which our modified model represents the true data generating process. We simulate from our model with parameter values set to the same values as above in Figure 3. We then study to what extent an analyst applying the (misspecified) standard Merton model would arrive at misleading conclusions about bank default risk.

Figure 5 shows the simulated actual risk-neutral default probabilities in our model (blue) and those estimated based on the Merton model (red) applied to our simulated data. The Merton model default probabilities are obtained by using the simulated equity values and instantaneous volatilities to extract asset values and asset volatilities under the (false) assumption of a log-normal asset value process (see Appendix A). This corresponds to the practice of applying the Merton
Figure 5: Risk-neutral default probabilities as function of current loan portfolio value: Merton (red) and modified (blue)

model to recent empirical volatilities and observed equity values. As the Figure shows, the Merton model underestimates the probability of default for moderate (and, in most cases, more realistic) default probabilities. In these good states of the world, the default probabilities are massively understated when the (misspecified) Merton model is applied. If the borrower asset values are relatively high, bank equity volatility is very low because bank asset volatility is very low. However, asset volatility could quickly rise if asset values fall, which means that it is much more likely than one would think based on the Merton model that the default threshold could be reached. Low instantaneous equity volatility hence does not mean that the bank operates at a high distance to default. However, an analyst applying the standard Merton model with constant asset volatility would miss these nonlinear risk dynamics. Within the standard Merton model, the analyst would interpret the low instantaneous equity volatility as a high distance to default and hence low default risk. Thus, particularly in good times, application of the Merton model is likely to lead to severe underestimation of bank default risk. Only in a severely distressed situation, when asset values are depressed and default is almost certain, the Merton model slightly overstates risk-neutral default
probabilities. Here the above effect works in reverse.

Figure 6 presents the implied credit spread of the bank’s 5-year debt. The credit spread reflects the product of the risk-neutral probability of default (as shown in Figure 5) and the loss given default (which equals one minus the recovery rate), shown on an annualized basis. Since the recovery rate in a Merton style model can often be quite high in cases where the asset value at maturity ends up only slightly below the face value of the debt, the implied credit spreads are much lower than the risk-neutral default probabilities. Our main focus is again on how an analyst using the standard Merton model would severely underestimate credit spreads. For example, when borrower asset values are around one, and hence substantially above the face value of their loans, the bank’s credit spread according to the Merton model should be about 10 basis points. In contrast, the true credit spread is close to 1% and hence almost an order of magnitude bigger.

Figure 7 shows how application of the standard Merton model would severely estimate the value of a government guarantee. For illustration, we suppose that there is a 50% risk-neutral probability that the government will fully bail out the debt holders (and absorb the entire loss given default) in the event of default. The value of the government guarantee then is 0.5 times the value of
the bank’s default option. To interpret the magnitudes in Figure 7, recall that the aggregate face value of the bank’s loans in our simulations is around 0.80. As the figure shows, application of the standard Merton model would severely underestimate the value of the guarantee. For aggregate borrower asset values around one, the value of the guarantee is about 0.0015, i.e., about 0.2% of the face value of the bank’s loan portfolio, while the actual value is almost ten times as big.

### III Empirical Calibration

To find out how much, quantitatively, the standard Merton model and our modified model differ in their predictions about default probabilities and risk dynamics, we now calibrate these models with empirical data on bank’s capital structures and equity volatility.

### III.A Data

Our sample covers all commercial banks listed in the Federal Reserve Bank of New York’s CRSP-FRB linked dataset from 2002-2012 that are also covered by Compustat. We obtain equity returns
Table I: Model Inputs

Our sample covers all commercial banks listed in the Federal Reserve Bank of New York’s CRSP-FRB linked dataset from 2002-2012 that are also covered by Compustat. Market equity and book value of assets are normalized by $D$, i.e., the numbers shown in the table are the market equity/book debt and the book assets/book debt ratios. Equity volatility is annualized and estimated from daily stock returns over one-year moving windows. Our risk-free interest rate proxy is the Federal Reserve Board’s 1-year treasury bill rate series.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Equity</td>
<td>0.15</td>
<td>0.08</td>
<td>0.03</td>
<td>0.10</td>
<td>0.15</td>
<td>0.21</td>
<td>0.30</td>
</tr>
<tr>
<td>Equity Volatility</td>
<td>0.44</td>
<td>0.24</td>
<td>0.17</td>
<td>0.27</td>
<td>0.35</td>
<td>0.52</td>
<td>1.03</td>
</tr>
<tr>
<td>Book value of assets</td>
<td>1.46</td>
<td>0.10</td>
<td>1.27</td>
<td>1.40</td>
<td>1.48</td>
<td>1.53</td>
<td>1.62</td>
</tr>
<tr>
<td>Risk-free Rate</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Observations 18,609

and market value of equity from CRSP and the accounting data from Compustat Annual files for banks. Since accounting data is available at annual frequency, we linearly extrapolate them over the year to generate accounting numbers for the intermediate months. In our base model, we consider the entire debt value of firm (including demand and time deposits) as $D$.\(^1\) We compute the bank’s payout ratio, $\gamma$, as the sum of cash dividend on common and preferred equity as a fraction of the book value of assets. Appendix C provides more details on the construction of the bank-level variables. Our risk-free interest rate proxy is the Federal Reserve Board’s 1-year treasury bill rate series.

Summary statistics of key inputs used for our estimation are provided in Table I. Market equity and book value of assets are normalized by $D$, i.e., the numbers shown in the table are the market equity/book debt and the book assets/book debt ratios. Equity volatility is computed (in annualized form) from daily bank stock returns over one-year moving windows.

### III.B Model calibration

For both the standard Merton model and our modified model, we set the maturity of bank debt to $H = 5$. It is well known that models with only diffusive shocks have no chance in delivering realistic

\(^1\)Note that this is different from several earlier papers that estimate distance-to-default based on KMV’s model that only includes short-term debt and half of long-term debt as the face value of $D$. A large proportion of a typical bank’s debt consists of deposits. Some of these deposits do not have any stated maturity (e.g., demand deposits), whereas some of them can be withdrawn anytime with small penalty (e.g., time deposits). Hence we include the bank’s entire debt in $D$. Naturally, this approach provides a higher estimate of the bank’s default rates as compared to the earlier approach.
default risk and credit spread predictions for short-term debt. At short horizons, the differences between our modified model and the standard model are also relatively small. The non-linearity in banks’ asset payoffs becomes more relevant as the probability distribution of borrower asset values spreads out with longer horizons. Longer horizons may also be relevant even for investors in short-term debt (or guarantors of short-term debt). Solvency problems may not be immediately apparent when bad shocks are realized. Deterioration in asset values may be hidden for a while, perhaps facilitated by regulatory forbearance, and short-term debt may be rolled over even if the bank is actually insolvent. By the time default happens, additional losses may have accumulated.

We calibrate the standard Merton model with an iterative approach. We start with an initial guess about the unobserved volatility of assets. Based on this guess, we can solve, every day, for the unobserved asset value. We use these daily asset values over backward-looking one-year window to update our guess for volatility and repeat until convergence. Appendix B provides more detail.

For our modified model, we have several additional parameters that we fix exogenously, as shown in Table II. We set the depreciation rate $\delta = 0.005$. We assume that the maturity of loans issued by banks is $T = 10$ and that borrower asset values have a pairwise correlation of $\rho = 0.5$. This correlation assumption is not crucial. A lower $\rho$ implies more smoothing of the kinks in the payoff function in Figure 2, but otherwise its effects are limited because any change in $\rho$ requires an offsetting change in $\sigma$ when we calibrate the model to match observed equity value and volatility. We set $\gamma$ equal to the observed payout rate of the bank.

To calibrate the modified model, we take $r$ to be the 10-year Treasury yield. Our model abstracts from a term structure of risk-free rates and we use the 10-year Treasury yield as an approximation. With normalize $D = 1$. The face value of the loans in the bank’s portfolio represents, roughly, the book assets of the bank. To take into account the zero-coupon nature of the loans, we adjust

### Table II: Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>Borrower Asset Depreciation Rate</td>
<td>0.005</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Bank payout Rate</td>
<td>0.002</td>
</tr>
<tr>
<td>$T$</td>
<td>Bank Loan Maturity</td>
<td>10 years</td>
</tr>
<tr>
<td>$H$</td>
<td>Bank Debt Maturity</td>
<td>5 years</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Borrower asset value correlation</td>
<td>0.5</td>
</tr>
<tr>
<td>$\ell$</td>
<td>Loan-to-Value Ratio</td>
<td>$0.8e^{(\mu-r)T}$</td>
</tr>
</tbody>
</table>
We calibrate the Merton model and our modified model twice a year (beginning of Q2 and Q4) from 2002-2012 based on the data summarized in Table I. For each bank in each calibration period, we compute the risk-neutral default probability from the two models. The table reports summary statistics for these risk-neutral default probabilities for the whole panel of banks over the full sample period.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merton Model RN prob(default)</td>
<td>0.08</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.82</td>
</tr>
<tr>
<td>Modified Model RN prob(default)</td>
<td>0.07</td>
<td>0.12</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.06</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Observations 18,609

the book value for the time until loan maturity and we set $F_1$ equal to the inverse of the observed book leverage times $\exp(7.5r)$. We then look for values of $dW_0$ and $\sigma$ that allow us to match the empirically observed equity value and equity volatility (over a 12-month backward looking window) of the bank with the model implied value.

### III.C Model-implied risk-neutral default probabilities

Table III gives the summary statistics of the two calibrated models’ implied risk-neutral default probabilities (RNPD) for our panel of banks. We calibrate each model twice a year (beginning of Q2 and Q4) from 2002-2012. The average RNPD is quite similar in our modified model compared with the standard Merton model. However, the Merton model RNPD is much more positively skewed. The reason is that when a bank is not in distress, the implied RNPD in the Merton model is very low because it is based on the assumption that the bank’s assets have constant volatility. In contrast, our model takes into account that bad shocks to asset values in the future would drive up the volatility of the bank’s assets (as borrower’s default options move into the money), which drastically shrinks the distance to default and hence raises the RNPD in times when asset values are relatively high.

Figure 8 further illustrates the different behavior of the RNPD from the two models in good times and bad times. The figure shows the average RNPD across all banks each quarter from 2002-2012. The modified model’s RNPD is two to three times as high as Merton RNPD during the time before the financial crisis. Expressed in terms of annualized credit spreads, these RNPD in 2006 would correspond, roughly, to 10 basis points for the Merton model and close to 100 basis
Figure 8: Comparison of calibrated risk-neutral default probabilities
Table IV: Differences in Model-Implied Risk-Neutral Default Probability: Comparison Between High- and Low-VIX Periods

The dependent variable is the log risk-neutral default probability from our modified model minus the log risk-neutral default probability from the standard Merton model for our panel of banks from 2002-2012. Explanatory variables include a dummy for quarters with below median VIX index, and the bank’s log market equity (normalized by to total debt, as in Table I).

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low VIX</td>
<td>2.426***</td>
<td>1.720***</td>
<td>5.594***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity</td>
<td>2.840***</td>
<td>2.257***</td>
<td>1.407***</td>
<td>-0.185</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.31)</td>
<td></td>
</tr>
<tr>
<td>Low VIX x Equity</td>
<td>1.905***</td>
<td>1.713***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.112***</td>
<td>8.177***</td>
<td>6.079***</td>
<td>4.188***</td>
<td>0.962**</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>YQ Fixed Effect</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>18608</td>
<td>18608</td>
<td>18608</td>
<td>18608</td>
<td>18608</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.351</td>
<td>0.370</td>
<td>0.404</td>
<td>0.421</td>
<td>0.460</td>
</tr>
<tr>
<td>Absorbed FE</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
</tr>
<tr>
<td>Clustered by</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
<td>permco</td>
</tr>
</tbody>
</table>

$p$-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

points for our modified model.

This behavior of the relative RNPDs is in line with the intuition that bank assets have nonlinear debt-like payoffs. In good times, the Merton model’s assumption of constant asset volatility produces very high distance to default and low RNPD. In our model, the analysis recognizes that the volatility of bank assets in good times (in the flat part of the concave asset payoff region in Figure 2) is low, but that it can quickly rise after a bad shock (when the bank gets into the downward sloping asset payoff region). Once the asset value has suffered a sufficiently big bad shock, the asset payoff is dominated by the linear downward sloping region and the kink is not playing much of a role. In this case, the predictions of the Merton model and our modified model are relatively similar. This is why after the onset of the financial crisis in 2007, the difference between the RNPD shrinks and eventually inverts in 2008.

We further illustrate this cyclical behavior of the RNPD differences between the Merton model
and our modified model with the following panel regression

\[
\log \frac{RNPD_{it}^{Modified}}{RNPD_{it}^{Merton}} = \alpha_i + \beta_1 \times LowVIX_t + \beta_2 \times \log E_{it} + \beta_3 \times LowVIX_t \times \log E_{it} + \epsilon_{it} \quad (20)
\]

The dependent variable is the log difference in default probabilities for a bank \(i\) in quarter \(t\). VIX is the CBOE index of implied volatilities on S&P500 index options. LowVIX equals one for quarters with below median VIX, zero otherwise. \(E\) measures each bank’s market equity (normalized by total debt, as in Table I). The regression results in Table IV show that the modified model RNPD is significantly higher during “quiet” periods, and for high equity banks. Column (IV) further shows the interaction effect: The standard Merton model delivers a lower default probability for banks with higher equity during low VIX periods.

Table V presents results of a similar regression model, but with the level of the S&P500 index during our sample period as the measure of good/bad times.

\[
\log \frac{RNPD_{it}^{Modified}}{RNPD_{it}^{Merton}} = \alpha_i + \beta_1 \times HighS&P_t + \beta_2 \times \log E_{it} + \beta_3 \times HighS&P_t \times \log E_{it} + \epsilon_{it} \quad (21)
\]

\(HighS&P\) equals one for quarters with above median level of S&P500 index, zero otherwise. The regression results show that the modified model RNPD is significantly higher during periods with high stock market valuation, and for high equity banks. Column (IV) further shows the interaction effect: The standard Merton model implies a lower default probability for banks with higher equity during periods in which the S&P500 index is high.

**IV Conclusion**

The standard assumption of a log-normally distributed firm asset value is not appropriate when applying structural models of default risk to banks. Banks’ assets are risky debt claims with capped upside and hence the asset payoff is nonlinear, with embedded optionality. As a consequence, bad shocks to bank asset values lead to a rise in asset volatility, unlike in the standard model where asset volatility is constant. A bad shock to asset values therefore reduces the distance to default much more than it would in the standard model. Our modification of the standard model takes
Table V: Differences in Model-Implied Risk-Neutral Default Probability: Comparison Between High- and Low-Stock Market Index Periods

The dependent variable is the log risk-neutral default probability from our modified model minus the log risk-neutral default probability from the standard Merton model for our panel of banks from 2002-2012. Explanatory variables include a dummy for quarters with below median S&P 500 Index and the bank’s log market equity (normalized by to total debt, as in Table I).

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>High S&amp;P Level</td>
<td>0.939***</td>
<td>0.647****</td>
<td>2.914***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Equity               |       | 2.840*** | 2.769*** | 2.323*** | -0.135
|                      |       | (0.00) | (0.00) | (0.00) | (0.45) |
| High S&P Level x Equity | 1.110*** |      |       | 1.346*** |
|                      | (0.00) |      |       | (0.00) |
| Constant             | 1.936*** | 8.177*** | 7.715*** | 6.783*** | 1.053*** |
|                      | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| YQ Fixed Effect      | No    | No    | No    | No    | Yes |
| Observations         | 18608 | 18608 | 18608 | 18608 | 18608 |
| $R^2$                | 0.287 | 0.370 | 0.375 | 0.381 | 0.457 |
| Absorbed FE          | permco | permco | permco | permco | permco |
| Clustered by         | permco | permco | permco | permco | permco |

*p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

this effect into account and leads to substantially different assessment of distance to default and bank risk dynamics. In good times, when asset values are high, the standard model would suggest risk-neutral default probabilities that are essentially zero, while a model that is consistent with the options-on-options nature of bank equity and debt suggests risk-neutral default probabilities that are substantially higher than zero. For the same reasons, the standard model would severely understate the value of implicit or explicit government guarantees. Furthermore, the standard model also underestimates the degree to which banks’ equity risk rises in response to a an adverse shock to asset values.

Our focus in this paper is on addressing the fundamental issue that a banks’ asset value cannot be log-normally distributed. As we have shown, this issue has first-order consequences for default risk evaluation and our modified structural model is useful to explore the economic drivers of bank default risk. Of course, a simple structural model of the kind we use here omits many additional features that would be necessary to realistically describe banks’ default risks. Extensions of the
model could explore jumps in asset values, default due to liquidity problems, complex maturity and seniority structures of banks’ debt, and various forms of explicit and implicit government support.
References


Crosbie, Peter J., and Jeffrey R. Bohn, 2001, Modeling default risk, KMV LLC.


Appendix

A Distance to default estimation in the standard Merton model: Simulations

The firm’s zero-coupon debt has face value $D$ and matures at $T$. The asset value $V$ evolves according to

$$
\frac{dV_t}{V_t} = (r - \gamma)dt + \sigma_v dB_t \tag{22}
$$

where $\gamma$ is the cash payout rate. The value of the firm’s equity then is

$$
S_t = C(V_t, D, r, \gamma, T - t, \sigma_v) + (1 - \exp[-\gamma(T - t)])V_t \tag{23}
$$

where $C(.)$ is the Black-Scholes call option price,

$$
C(V_t, D, r, \gamma, T - t, \sigma_v) = V_t \exp[-\gamma(T - t)]N(d_1) - D \exp[-r(T - t)]N(d_2) \tag{24}
$$

and

$$
d_1 = \frac{\log V_t - \log D + (r - \gamma + \sigma_v^2/2)}{\sigma_v \sqrt{T - t}} \\
d_2 = d_1 - \sigma_v \sqrt{T - t},
$$

while equity volatility follows from the leverage ratio of the call option replicating portfolio as

$$
\sigma_{s,t} = \frac{V_t \exp[-\gamma(T - t)]N(d_1)}{S_t} \sigma_v. \tag{25}
$$

In our simulations in Section II.A where we apply the Merton model (as a misspecified model) to data generated from our modified model, we use the simulated values of $S_t$ and the instantaneous equity volatility $\sigma_{s,t}$ to solve equations (23) and (25) for $V_t$ and $\sigma_v$. Based on $V_t$ and $\sigma_v$ we can then compute the risk-neutral (RN) distance-to-default as

$$
DD_{BSM} = \frac{\log V_t - \log D + (r - \gamma - \sigma_v^2/2)}{\sigma_v \sqrt{T - t}}
$$

The corresponding implied RN default probability, also called the expected default frequency (EDF), can be computed as follows:

$$
EDF_{BSM} = \Phi \left( \frac{-\log V_t + \log D - (r - \gamma - \sigma_v^2/2)}{\sigma_v \sqrt{T - t}} \right)
$$

where $\Phi(.)$ is the standard normal CDF.
B Distance to default estimation in the standard Merton model: Empirical data

There are two approaches to empirically implement this model. In the first approach, one can take empirically observed equity values and equity volatility estimates to solve equations (23) and (25) simultaneously, as we do in our simulations. However, in practice, market leverage changes too frequently to provide reasonable estimates of asset volatility via equation (25). The second approach involves an iterative process to estimate the asset value and asset volatilities from equity market data. This approach has been adopted by previous researchers such as Crosbie and Bohn (2001), ?, and Bharath and Shumway (2008). We adopt a similar iterative approach for our estimation exercise.

We start with an initial guess for \( \sigma_v \). To obtain this initial guess, we add up the market value of equity and book value of debt of the bank on a daily basis. Based on these daily approximate asset values, we compute daily log asset returns and then take their standard deviation as our initial guess for \( \sigma_v \). With this value of \( \sigma_v \) we compute the \( V_t \) using equation (23) and the observed \( S_t \) each day during the previous year. We then re-compute the implied log return on assets each day based on these estimates. With the estimated log return values, we compute new estimates of \( \sigma_v \) in each iteration. We iterate on \( \sigma_v \) in this manner till it converges. The convergence is achieved when the absolute difference between two successive \( \sigma_v \) estimates is less than 0.001. With this final estimate of \( \sigma_v \), we solve for the final estimate of \( V_t \).

C Data construction

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
<th>Construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Market Equity Value of Bank</td>
<td>CRSP shrc x shrout</td>
<td></td>
</tr>
<tr>
<td>sE</td>
<td>Stock Return Volatility</td>
<td>CRSP</td>
<td>Stock return volatility over the past year</td>
</tr>
<tr>
<td>F</td>
<td>Face Value of Borrower’s Loan</td>
<td>Compustat (BV of bank debt + equity + loan loss reserve) ( \times e^{-rf \times 10} )</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>Risk-free rate</td>
<td>FRB</td>
<td>log 1-year risk-free rate</td>
</tr>
<tr>
<td>D</td>
<td>Face Value of Bank Debt</td>
<td>Compustat</td>
<td>short-term debt + long-term debt + deposits ( (dlc+dltt+dptc) )</td>
</tr>
</tbody>
</table>

Notes:
All book values are obtained from Annual Compustat Files for Banks.
Book values are linearly interpolated between the annual year end dates.
All book values are obtained from Annual Compustat Files.
Stock return volatility is based on daily returns.