



Research Paper

Dynamic Loan Losses: Estimation and Implications

Date: 2006

Reference Number: 6/3

DYNAMIC LOAN LOSS DISTRIBUTIONS: ESTIMATION AND IMPLICATIONS

Robert Lamb
Imperial College
London

William Perraudin
Imperial College
London

First version: August 2005

This version: August 2006*

Abstract

This paper develops new dynamic versions of Vasicek's widely used single factor, loan loss distribution, and employs them to analyze capital and risk in loan portfolios. We show that ignoring loan loss dynamics leads to a systematic upward bias in capital calculations. We calculate the impact of this bias for aggregate US loan loss rates and compare our results with the calibration implicit in the Basel II capital charges. Backing out innovations from different loan loss series, we study the correlations between six different loan sectors: corporate, credit card, other retail, agricultural, leases and mortgages.

*The authors' may be contacted at Tanaka Business School, Imperial College, London SW7 2AZ or at rlamb@imperial.ac.uk or wperraudin@imperial.ac.uk.

1 Introduction

Estimates of the distribution of losses on credit portfolios are important inputs to capital decisions by lenders and to the design of financial regulations by supervisors. The most widely used model of loan losses is that suggested by Vasicek (1991). Under this approach, defaults by individual obligors are driven by realizations of latent random variables. If the latent variable for an exposure falls below a given threshold, the obligor defaults. Dependence between defaults by different obligors is introduced by allowing the corresponding latent variables to be correlated. As the number of exposures in the portfolio increases to infinity, Vasicek shows that the distribution of the loss rate (the fraction of obligors that default) converges to a simple closed form expression.

The task of calibrating the loss rate models reduces to that of estimating the correlation of the latent variables. To see why this is true, note that as the number of credit exposures in a portfolio increases, if defaults on the individual exposures are independent, the fraction of defaults will converge to a constant. Conversely, if the default rate fluctuates over time, it must reflect correlation in the underlying latent variables. Starting with measures of the time series variability of aggregate portfolio default rates, one may therefore deduce correlation parameters for the underlying latent variables.

Vasicek's loan loss distribution has been widely employed by academic and industry researchers in modeling credit portfolio losses. It has also served as the basis for the capital charge formulae or "capital curves" contained in the Basel II proposals. Any bank regulated under Basel II will calculate the regulatory capital it must hold against a particular credit exposure by plugging variables describing the exposure (such as default probability and loss given default) into the relevant formula.

The problem with basing risk measures and capital calculations on the simple Vasicek model, however, is that it is a purely static model in which loan losses are realized randomly and unpredictably over time. In practice, default rates on portfolios move in a very predictable way from period to period. The probability of a high (low) default rate in year t is much greater if the default rate was high (low) in $t - 1$. But this implies that the risk faced by a bank in year t is much less if it knows what the default rate was the period before.

Unfortunately, a common practice in benchmarking the Vasicek model from data

is to ignore the time series properties of loan losses, in particular their predictability, and to deduce correlation parameters from the unconditional distribution of aggregate annual loss rates. This approach exaggerates the risk faced by lenders.

The contributions of this paper are

1. To generalize the Vasicek model to allow for autocorrelated loan losses. We formulate models in which the underlying factors driving the latent variables are autoregressive processes and we show how the transformed loss rates inherit this time series structure.
2. To implement the model empirically using loan loss data for US banks. We apply a Maximum Likelihood approach to aggregate portfolio loss rates to deduce parameters of the latent variable factors.
3. To compare the implied capital with the regulatory capital charges implied by the Basel II proposals. The Basel capital formulae are directly based on the single factor model proposed by Vasicek. Our model nests the Vasicek model so we can show precisely how capital calculations are affected by correctly incorporating autocorrelation in the loss rates. The calibration and estimation exercises we perform also demonstrate that even leaving the autoregressive properties of loss rates out of account, in the case of retail portfolios, the Basel II parameterization is much more conservative than is justified by the data.
4. To back out the innovations in losses for the six different loan categories we study and to examine the correlation structure. This sheds light on the degree to which banks may diversify by extending loans to different credit sectors. Again, we compare our results with the Basel capital charges.

Our study is related to the recent empirical literature on modeling loss rate and ratings transitions. Several papers within this literature employ static one-period Vasicek models to investigate loan loss data and to analyze the capital implications for financial firms. Perli and Neyda (2004) compare several different portfolio credit risk models designed for retail portfolios and compare the results obtained with the capital charges implied by the current Basel II proposals. de Andrade and Thomas (2004) use structural and reduced form credit risk models to investigate Brazilian credit bureau scores. Dietsch and Petey (2004) examine whether Small and Medium Enterprise (SME) credit exposures resemble more retail or corporate exposures using

French and German SME data by estimating asset correlations using the Vasicek distribution. Cowan and Cowan (2004) calculate default correlations for sub-prime residential mortgages. None of these papers allows for the time series behavior of losses in their statistical modeling or the implications they draw for risk and capital.

Another relevant literature concerns the cyclical pattern of loss rates and ratings transition probabilities. Early studies include Wilson (1997a), Wilson (1997b), and Nickell, Perraudin, and Varotto (2000). Allen and Saunders (2003) provide a survey. Recent contributions include Bangia, Diebold, Kronimus, Schagen, and Schuermann (2002), Couderc, Renault, and Scaillet (2003), Gorton and He (2003), Koopman and Lucas (2005), Koopman, Lucas, and Klaassen (2005), McNeil and Wendin (2005b), and McNeil and Wendin (2005a). As one might expect, this literature more directly examines the time series nature of loss and transition rates.

More broadly, our study is related to the substantial and growing literature on modeling correlations in credit portfolios. Embrechts, Lindskog, and McNeil (2003), Frey and McNeil (2003), and Kiesel and Schmidt (2004) set out general results. Applications to CDO valuation include Li (2000), Duffie and Garleanu (2001), Hull and White (2004), Laurent and Gregory (2003), the studies surveyed in Burtschell, Gregory, and Laurent (2005), Sidenius, Piterbarg, and Anderson (2004) and Schonbucher (2005).

Of these papers, many focus on one-period loss distributions, i.e., simple static models (see Embrechts, Lindskog, and McNeil (2003), Frey and McNeil (2003), Schonbucher (2004), Schonbucher (2002) and Wehrspohn (2003), for example). Where a fully dynamic model is required (in order, say, to price a CDO with a potentially complex cash flow waterfall), a joint distribution of the times to default is used to simulate dependent default times (see Li (2000) and subsequent papers that have followed his approach).

Hazard rate models for pricing structured products such as Duffie and Garleanu (2001) possess fully specified dynamic structure. In a recent contribution, Das, Duffie, Kapadia, and Saita (2005) investigate the ability of simple hazard rate models to explain correlations in equity-based default probabilities. Hanson, Pesaran, and Schuermann (2006) examine the time series structure of equity factors within structural credit risk models, relating the dynamics to conditioning macro variables.

In the last section of the paper, we back out the innovations in loan losses for different sectors and study their correlations. This is an important topic since little

is known empirically about the how important diversification effects are across the different lending divisions of banks. Recent papers have begun to study correlations in credit portfolios. Heitfield and Chomsisengphet (2005) look at sector correlations in equity factors used in typical credit portfolio models. Varotto (2005) looks at sector and country correlations in bond spreads.

Finally, our paper may be compared to other academic research that has commented on features of the Basel II parameterization. Examples include Altman and Saunders (2001), Kupiec (2001) and Perli and Neyda (2004).

2 Loan Loss Models

2.1 Loan Losses with Autocorrelation and Trends

Suppose there are n obligors. If he is not in default at $t - 1$, the i th obligor defaults at t if a latent variable, $Z_{i,t}$ satisfies $Z_{i,t} < -c$ for a constant c .

Suppose that the $Z_{i,t}$ for $t = 0, 1, 2, \dots$ and $i = 1, 2, \dots, n$ satisfy a factor structure in that:

$$Z_{i,t} = \sqrt{\rho}X_t + \sqrt{1 - \rho}\epsilon_{i,t} . \quad (1)$$

Assume that:

$$X_t = \sqrt{\beta}X_{t-1} + \sqrt{\lambda}\eta_t \quad (2)$$

where $\epsilon_{i,t}$ and η_t are standard normal and independent for pairs of obligors i and j and for all dates t .

Typically, when latent variable models such as that described by equation 1 are formulated, the shocks are assumed to have unit variance. To preserve comparability with other models, we construct our model so that X_t has unit unconditional variance. Since

$$X_t = \sum_{i=0}^{\infty} \left(\sqrt{\beta}\right)^i \sqrt{\lambda}\eta_{t-i} \quad (3)$$

$$\text{Variance}(X_t) = \sum_{i=0}^{\infty} \beta^i \lambda = \frac{\lambda}{1 - \beta} \quad (4)$$

if we choose $\lambda = 1 - \beta$, the unconditional variance of X_t is unity and

$$X_t = \sqrt{\beta}X_{t-1} + \sqrt{1 - \beta}\eta_t . \quad (5)$$

If X_t has unit variance unconditionally, the $Z_{i,t}$ has a standard normal unconditional distribution. The unconditional probability of default for the i th obligor, q , then satisfies:

$$\Phi^{-1}(q) = -c. \quad (6)$$

Now, consider the probability of default for i in our model conditional on information at $t-1$. Default occurs when:

$$\sqrt{\rho}X_t + \sqrt{1-\rho}\epsilon_{i,t} < -c \quad (7)$$

$$\sqrt{\rho}\sqrt{1-\beta}\eta_t + \sqrt{1-\rho}\epsilon_{i,t} < -c - \sqrt{\rho}\sqrt{\beta}X_{t-1}. \quad (8)$$

The left hand side of equation (8) is distributed as $N(0, 1-\rho\beta)$. So the default probability conditional on information at $t-1$ is:

$$p_{i,t} = \Phi\left(\frac{-c - \sqrt{\rho}\sqrt{\beta}X_{t-1}}{\sqrt{1-\rho\beta}}\right). \quad (9)$$

Conditional on the common factor η_t and on X_{t-1} , defaults are independent across individual obligors. So, denoting $P(k, n)$ as the probability of observing k defaults out of n obligors conditional on X_{t-1} and removing the conditioning of the shock, η_t , by integrating over its support, gives:

$$P(k, n) = \binom{n}{k} \int_{-\infty}^{\infty} \Phi\left(\frac{-c - \sqrt{\rho}\sqrt{\beta}X_{t-1} - \sqrt{\rho}\sqrt{1-\beta}\eta_t}{\sqrt{1-\rho}}\right)^k \times \left[1 - \Phi\left(\frac{-c - \sqrt{\rho}\sqrt{\beta}X_{t-1} - \sqrt{\rho}\sqrt{1-\beta}\eta_t}{\sqrt{1-\rho}}\right)\right]^{n-k} d\Phi(\eta_t). \quad (10)$$

Adopting the change of variables:

$$s(\eta) \equiv \Phi\left(\frac{-c - \sqrt{\rho}\sqrt{\beta}X_{t-1} - \sqrt{\rho}\sqrt{1-\beta}\eta}{\sqrt{1-\rho}}\right). \quad (11)$$

Then:

$$P(k, n) = - \binom{n}{k} \int_0^1 s^k (1-s)^{n-k} d\Phi\left(\frac{-(\sqrt{1-\rho}\Phi^{-1}(s) + c + \sqrt{\rho}\sqrt{\beta}X_{t-1})}{\sqrt{\rho}\sqrt{1-\beta}}\right) \quad (12)$$

But:

$$-d\Phi(f(s)) = d\Phi(-f(s)) \quad (13)$$

so

$$P(k, n) = \binom{n}{k} \int_{-\infty}^{\infty} s^k (1-s)^{n-k} dW(s), \quad (14)$$

where

$$W(s) \equiv \Phi \left(\frac{\sqrt{1-\rho}\Phi^{-1}(s) + c + \sqrt{\rho}\sqrt{\beta}X_{t-1}}{\sqrt{\rho}\sqrt{1-\beta}} \right) \quad (15)$$

Now, consider what happens as $n \rightarrow \infty$. Let θ denote the fraction of the pool that defaults. Then:

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor n\theta \rfloor} P(i, n) = \int_0^1 \left(\lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor n\theta \rfloor} \binom{n}{i} s^i (1-s)^{n-i} \right) dW(s) \quad (16)$$

$$= \int_0^1 1(s < \theta) dW(s) \quad (17)$$

$$= W(\theta) - W(0) = W(\theta). \quad (18)$$

Hence, the loss distribution conditional on X_{t-1} is:

$$W(\theta_t) \equiv \Phi \left(\frac{\sqrt{1-\rho}\Phi^{-1}(\theta_t) - \Phi^{-1}(q) + \sqrt{\rho}\sqrt{\beta}X_{t-1}}{\sqrt{\rho}\sqrt{1-\beta}} \right). \quad (19)$$

So the transformed loss rate $\tilde{\theta}_t \equiv \Phi^{-1}(\theta_t)$ is Gaussian and satisfies:

$$\tilde{\theta}_t \equiv \Phi^{-1}(\theta_t) \sim N \left(\frac{1}{\sqrt{1-\rho}} \left[\Phi^{-1}(q) - \sqrt{\rho}\sqrt{\beta}X_{t-1} \right], \frac{\rho(1-\beta)}{1-\rho} \right). \quad (20)$$

One may express this as:

$$\tilde{\theta}_t = \frac{1}{\sqrt{1-\rho}} \left[\Phi^{-1}(q) - \sqrt{\rho}\sqrt{\beta}X_{t-1} \right] - \frac{\sqrt{\rho}\sqrt{1-\beta}}{\sqrt{1-\rho}} \eta_t \quad (21)$$

where η_t is the shock in equation (2) above. Note that the negative sign before the last term enters because of our change of variable $-d\Phi(\eta(s)) = d\Phi(-\eta(s))$.

Solving (21) for X_{t-1} , we obtain:

$$X_{t-1} = \frac{1}{\sqrt{\rho}\sqrt{\beta}} \left[\Phi^{-1}(q) - \sqrt{1-\rho}\tilde{\theta}_t - \sqrt{\rho}\sqrt{1-\beta}\eta_t \right]. \quad (22)$$

Lagging this equation and substituting in

$$X_{t-1} = \sqrt{\beta}X_{t-2} + \sqrt{1-\beta}\eta_{t-1} \quad (23)$$

yields the following result:

Proposition 1 *Suppose that an individual obligor who is not in default at $t-1$ defaults at t if a latent variable $Z_{i,t} < -c$. Assume that the $Z_{i,t}$ satisfy the Gaussian*

autoregressive factor structure given by equations (1) and (2). Let θ_t denote the loss rate, i.e., the fraction of obligors that default. As $n \rightarrow \infty$, the transformed loss rate, $\tilde{\theta}_t \equiv \Phi^{-1}(\theta_t)$ converges to the following Gaussian order-1 autoregressive process:

$$\tilde{\theta}_t \equiv \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q) + \sqrt{\beta} \left[\tilde{\theta}_{t-1} - \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q) \right] - \frac{\sqrt{\rho}\sqrt{1-\beta}}{\sqrt{1-\rho}}\eta_t. \quad (24)$$

Hence, the transformed loss rate at t conforms to the following Gaussian distribution:

$$\tilde{\theta}_t \equiv \Phi^{-1}(\theta_t) \sim N \left(\frac{(1-\sqrt{\beta})}{\sqrt{1-\rho}}\Phi^{-1}(q) + \sqrt{\beta}\tilde{\theta}_{t-1}, \frac{\rho(1-\beta)}{1-\rho} \right). \quad (25)$$

Now, suppose that the cut-off point for default evolves deterministically over time in that an obligor not in default at $t-1$ defaults at t if $Z_{i,t} < -c(t)$ for a function $c(\cdot)$. Define $q_t \equiv \Phi(-c(t))$. Then, one may easily show that $\tilde{\theta}_t$ follows the process

$$\tilde{\theta}_t \equiv \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q_t) + \sqrt{\beta} \left[\tilde{\theta}_{t-1} - \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q_t) \right] - \frac{\sqrt{\rho}\sqrt{1-\beta}}{\sqrt{1-\rho}}\eta_t. \quad (26)$$

This process is autoregressive around a non-linear trend.

2.2 The General AR Case

The above analysis may be generalized to allow for any autoregressive process for the factor, X_t . Suppose that:

$$X_t = \sqrt{\alpha}B(L)X_{t-1} + \sqrt{1-\alpha}\eta_t \quad (27)$$

where

$$B(L) X_{t-1} \equiv \sum_{i=1}^M \xi_i X_{t-i}. \quad (28)$$

Choose $\sqrt{\alpha}$, ξ_i $i = 1, 2, \dots, M$ so that the unconditional variance of X_t is unity.

Similar arguments to those used before imply:

$$\tilde{\theta}_t = \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q_t) - \sqrt{\rho}\sqrt{\alpha}B(L)X_{t-1} - \frac{\sqrt{\rho}\sqrt{\alpha}}{\sqrt{1-\rho}}\eta_t. \quad (29)$$

If the roots of the polynomial lag operator $B(L)$ lie outside the unit circle, then:

$$X_{t-1} = B^{-1}(L) \left(\frac{1}{\sqrt{\rho}\sqrt{\alpha}} \left[\Phi^{-1}(q_t) - \sqrt{1-\rho}\tilde{\theta}_t - \sqrt{\rho}\sqrt{1-\alpha}\eta_t \right] \right) \quad (30)$$

Lagging this equation by one period, substituting for X_{t-1} and X_{t-2} in:

$$X_{t-1} = \sqrt{\alpha}B(L)X_{t-2} + \sqrt{1-\alpha}\eta_{t-1}, \quad (31)$$

and then rearranging, one obtains:

Proposition 2 *When X_t follows the general autoregressive process in equation (27)*

$$\tilde{\theta}_t = \frac{1}{\sqrt{1-\rho}}\Phi^{-1}(q_t) + \sqrt{\alpha}B(L)\left[\tilde{\theta}_{t-1} - \frac{\Phi^{-1}(q_{t-1})}{\sqrt{1-\rho}}\right] - \frac{\sqrt{\rho}}{\sqrt{1-\rho}}\sqrt{1-\alpha}\eta_t. \quad (32)$$

3 Estimation

3.1 Data

To implement our model, we employ quarterly aggregate charge-off data for all US banks for the period 1973 to 2005. This is published by the Federal Reserve Board. Note, that the data is presented net of recoveries. We therefore scale each of the loss series by dividing through by the LGD for each series. The assumed LGDs were taken from the Basel Committee on Banking Supervision (1999) and can be found in Table 2.

A more standard approach would possibly be to provide estimations using the same series but in an annualized form. However, we would only have twenty annual observations, so the statistical reliability of results obtained using this data would then be an issue. We therefore report results for the quarterly series.

Note that some past studies have focussed on arrears rates. See, for example, Whitley, Windram, and Cox (2004) who employ APACS arrears data and Black and Morgan (1998) and Dunn and Kim (1999) who use survey data from the Federal Reserve Board's Survey of Consumer Finances. We prefer to focus on bank charge-offs because these more closely resemble defaults as perceived by the banks.

There may be problems with charge-off rates. 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. One may hope that the fact that we employ data that aggregates charge-offs from many banks will mitigate this problem.

Plots of the six charge-off rate series are shown in Figure 1. It is apparent that the credit card charge-off rate has a long-run trend around which it has fluctuated

historically. Other series appear to cycle but exhibit less evidence of a trend. In the exercises we report below, we de-trend the credit card series before performing other statistical analysis. This is a simple alternative to freely estimating a loss rate model with both a trend and an autoregressive component.

In general, visual inspection of the series seems to confirm the presence of trends and cyclical behavior in loss rates, reinforcing the basic point of this paper that the unconditional and conditional distributions of charge-offs are very different.

3.2 Estimating the Model

To estimate the simple static Vasicek loss distribution, one may follow one of two approaches. The first of these has been used to our knowledge in the industry and by regulators and consists of solving a simple moment condition for the implicit correlation parameter. To be specific, if θ is the fraction of loans that default in an infinitely granular pool satisfying the Vasicek assumptions, one may show that:

$$E(\theta^2) = \Phi_2(\Phi^{-1}(E(\theta)), \Phi^{-1}(E(\theta)), \rho) . \quad (33)$$

Here, $\Phi_2(x_1, x_2, \rho)$ is the joint distribution function for two standard normals, x_1 and x_2 , with a correlation parameter ρ . Given estimates of the first and second moments of aggregate losses from a sample, one may estimate the correlation parameter by inverting equation (33) to obtain ρ .

The second approach one may take uses the fact that under the Vasicek assumptions, the transformed loss rate, $\tilde{\theta}_t$ is normally distributed with a standard deviation that depends in a simple way on the asset correlation. See equation (25) for the dynamic distribution case. So regressing $\tilde{\theta}_t \equiv \Phi^{-1}(\theta_t)$ on a constant, one may infer the unconditional default probability from the coefficient on the constant and the correlation parameter ρ from the standard error of the regression residuals.

In this paper, we follow the second approach. It is easy to extend this to cases in which the transformed loss rate is autoregressive simply by regressing on lagged $\tilde{\theta}_t$ as well as on a constant. As we saw in equation (25), the transformed loss rate is conditionally Gaussian, and, in the AR(1) case, one may infer the unconditional default probability q , the square of the factor loading on its lagged level, β , and the correlation parameter ρ from the regression constant, the regression coefficient on $\tilde{\theta}_{t-1}$ and the standard error of the regression residuals.

In the past, we have estimated correlation parameters using both methods and obtained almost identical estimates.

3.3 Results

The estimation results for the various quarterly series are contained in Table 1. The first two rows show the unconditional mean and the unconditional standard deviations of loss rates. Credit cards exhibit a much lower credit quality than the other loan categories in that the mean loss is 2.27%. By contrast, agricultural loans, for example, exhibit a loss rate of just 0.39%. But the volatility of the credit card loss rate (at 0.34%) is of a similar order of magnitude to the other categories. This initially suggests that credit card loans have a more risky nature compared to the other loan types. But, the main influence in the riskiness of a loan type (i.e. extreme events such as all the credit holder default at once) is based upon the correlation parameter, ρ .

Transforming the loss rate series by applying $\Phi^{-1}(\cdot)$ reveals that the credit card transformed loss rate volatility is less than those for other series. This then translates into a smaller correlation parameter for credit cards, 0.39%, compared to 12.86% for agricultural and 7.32% for real estate. Other retail has a correlation parameter of 1.13% while corporate loans and leases (that are primarily corporate) have correlation parameters of 4.79% and 6.05%. This shows that series that are of a similar nature seem to display similar characteristics, a point that will be returned to in the common factor decomposition section.

So, the mapping between loss rates, their volatilities and the correlation parameter is complex and one cannot just judge the riskiness of a loan type based on the standard loss series. It is from the volatility of the transformed series that the factor correlation can be extracted and this parameter is the major influence upon rare events occurring.

In the lower part of Table 1, results are reported for models that include autocorrelation. Introducing autocorrelation substantially reduces the standard errors of the transformed loss rates. For example, it falls from 30.72% to 12.41% in the case of All Real Estate loans. Perhaps surprisingly since they reflect the variability in loss rates, the correlation parameters ρ are broadly comparable to those from the model without autocorrelation. This does not mean, however, that the risk characteristics of the loss rates is the same as the series are now mean-reverting.

To appreciate the risk implications of mean reversion, note that if some variable Y_t is autoregressive of order 1 (AR(1)) then, omitting shocks for simplicity it may be written as: $Y_t = c + bY_{t-1}$. So,

$$Y_t - Y_{t-1} = c + (b - 1)Y_{t-1} = (1 - b)(c/(1 - b) - Y_{t-1}) \quad (34)$$

Suppose that $(1 - b) > 0$. Then if $Y_{t-1} < c/(1 - b)$, $Y_t - Y_{t-1}$ will be positive and otherwise it will be negative. Thus, Y_t will gradually revert to the long run level of $c/(1 - b)$. A high value of $(1 - b)$ implies rapid reversion of Y_t to its long run level while low $1 - b$ implies slow reversion.

Note that the volatility of innovations in the factor equals $\sqrt{1 - b^2}$. Increasing b decreases risk by reducing the volatility of innovations but slows the pace of reversion. Long-term risk as measured by the unconditional volatility is constant because the lower volatility of innovations is just offset by slower mean reversion.

To explain the risk associated with this autocorrelation model first note that the unconditional volatility of the process has been fixed at unity. With an autocorrelated model, at any given time, the conditional volatility will be lower than the unconditional case due to the amalgamation of the risky shocks into the autocorrelation term. The conditional case is therefore a weighted sum of the previously realised shocks and an idiosyncratic term. By construction, the idiosyncratic term has a volatility being equal to $\sqrt{1 - b^2}$ and so its volatility is less than one. As this is the only risk in the conditional process, its risk will be lower than the unconditional case.

Two limiting cases can help to explain the overall risk occurring with the autocorrelated process. If $b = 0$, then the autocorrelation process reduces to the static one. In this case, the conditional risk will be equal to the unconditional risk. If, $b = 1 - \epsilon$ for a small $\epsilon > 0$, then there is nearly no conditional volatility present. Nearly all of the weight will be on the autocorrelation term and as described above the near zero weight on the idiosyncratic risk term is the only driving factor for the volatility of the process.

Over a short horizon, the value of b is also less important for risk. Indeed, over a 1-period horizon, risk is independent of b even though over a longer horizon it reduces it.

4 Capital Implications

4.1 Implied Capital

Recall that the Value at Risk for a portfolio held over a given horizon and with a confidence level α is defined to be the loss that is exceeded on a fraction α of occasions. Suppose that the bank that holds the credit card book sets its capital equal to the Value at Risk of its portfolio as a whole.

Suppose that losses on the bank's portfolio in period t are a function only of the single risk factor described above, namely X_t . The Marginal Value at Risk (MVaR $_{\alpha}$) for a single loan may be calculated as the expected loss on the exposure conditional on X_t being at its α -quantile. But the expected loss is just equal to the probability of default multiplied by the loss given default (LGD), i.e.,

$$\text{MVaR}_{\alpha} = \text{LGD} \times \Phi \left(\frac{\Phi^{-1}(p) - \sqrt{\rho}\sqrt{\beta}X_{t-1} - \sqrt{\rho}\sqrt{1-\beta}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right). \quad (35)$$

But the probability of default conditional on information at $t-1$ is given by equation (9). Substituting yields the proposition:

Proposition 3 *Under the above assumptions, the Marginal Value at Risk for a single loan, denoted MVaR $_{\alpha}$, is:*

$$\text{MVaR}_{\alpha} = \text{LGD} \times \Phi \left(\sqrt{1-\rho}\beta \frac{\Phi^{-1}(p_t) - \sqrt{\rho}\sqrt{1-\beta}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right). \quad (36)$$

4.2 Basel II Capital

The Basel II capital formulae are based on Unexpected Loss i.e., the MVaR $_{\alpha}$ minus the expected loss:

$$\text{Basel Capital Formula} = \text{LGD} \times \Phi \left(\frac{\Phi^{-1}(p) + \sqrt{\rho}\sqrt{1-\beta}\Phi^{-1}(\alpha^*)}{\sqrt{1-\rho}} \right) - \text{LGD} \times p. \quad (37)$$

where α^* equals $1-\alpha$ in our notation. Note here that the Basel formula is simplified by the absence of the expression $\sqrt{1-\rho}\beta$ and depends on p rather than the conditional default probability p_t . When $\beta = 0$ so the factor is no longer autocorrelated, the latter term simplifies to unity and $p_t = p$ for all t . Since $\phi^{-1}(\alpha^*) = -\phi^{-1}(\alpha)$, in this case, the Basel capital formula equals the MVaR $_{\alpha}$ expression in equation (36).

4.3 Regulatory Implications

It is a straightforward exercise to relate capital charges as used by regulators and banks if the data period is one year. We therefore convert the data from quarterly series to their annual representations for the following capital calculations.

We calculate capital under the Basel II rules and based on MVaRs implied by the estimated processes with and without autocorrelation. The Basel II calculation assumes a point-in-time default probability as a parameter for their formulae. We take the loss rate at the end of the sample period for this parameter as this can be viewed as a proxy for the conditional default probability.

Table 2 reports the capital requirements calculated by all models. The top half of the table displays the parameters assumed for the Basel II calculation. The lower half of the table reports the capital for the three different cases. It is immediately apparent that the capital implied by the autocorrelation case is less than the static case for all series. For the Real Estate category, the implied capital for the autocorrelation case is 0.28%, this is approximately 5 times less than the static case at 1.35%. Other series are closer however. The credit card calculation implies 2.40% in the static case and 2.22% in the autocorrelation case.

The Basel II capital comparisons are even more striking. In all cases with or without autocorrelation the implied capital is substantially less than what Basel suggests. As an example, the other consumer category implies 0.90% using the autocorrelation model, 1.79% with the static model and 8.07% under the Basel II rules. In all cases the order of magnitude is at least four times greater for the Basel II calculation than the autocorrelation model and in some cases the difference is even greater.

This is not surprising in the case of long-lived assets like real estate, leases and C&I as the Basel parameterization in these cases is based on economic-loss-mode MVaRs from ratings-based models like Creditmetrics. (Default mode capital curve formulae are just convenient functions that happens to have the right shape.) But for short term assets including retail (especially credit cards), this difference in the implied capital can be viewed as very important for banks and financial institutions that hold such books.

5 Factors and Correlations

5.1 Factors and Factor Shocks

From each transformed loss rate time series, one may use the equation

$$\tilde{\theta}_t = \frac{\Phi^{-1}(q) - \sqrt{\rho}X_t}{\sqrt{1 - \rho}} . \quad (38)$$

to infer the underlying factor value in each period. We calculate the factors driving both the static and autocorrelation models for each loan category. In the autocorrelation case, we deduce the common factor innovations:¹

$$\eta_t = \frac{X_t - \sqrt{\beta}X_{t-1}}{\sqrt{1 - \beta}} . \quad (39)$$

Figure 2 shows the factor innovations over time. If the autocorrelation model is correct then the first order lag should remove all time dependencies from the data and the shocks should follow a white noise process. Inspection suggests that there may be some clustering of shocks in Leases and Agricultural loans and possibly some higher order autocorrelation in Credit Card receivables.

As a specification test, we perform normality tests on the innovations from the different models (see Table 3). These include

1. The Lilliefors test (a generalization of the Kolmogorov-Smirnov test). For each potential value x , the Lilliefors test compares the proportion of values less than x with the forecast number implied by the normal distribution. The maximum difference, over all x , values is its test statistic. The asymptotic 95% quantile for rejection of normality is: 0.0978.
2. The Jarque-Bera test. This assesses whether the sample skewness and kurtosis differ to a statistically significant degree from their values under normality. We generate confidence levels for the Jarque-Bera test using Monte Carlos rather than using the asymptotically valid confidence level. The Monte Carlo confidence level for rejection of normality is 5.99.

The Lilliefors test suggests rejection of normality in that four of the six sets of innovations fail the test at the 95% level. However, when we perform the Jarque-Bera test, we find no rejections using the Monte-Carlo generated confidence level of 5.99.

¹An initial value of $X_0 = 0$ was assumed for each factor series.

5.2 Factor Correlations

A key question is what are the correlations between the factors of the different categories of loan losses? The answer can help shed light on the degree to which diversifying across different types of business will help a bank reduce its risk.

For the static model we calculate the correlation matrix of the factor series and for the autocorrelation model we calculate the correlation matrix of the factor innovations. Table 4 reports these matrices for quarterly data. (To save space, we do not exhibit the correlation matrices for annual data that we also estimated.)

The correlation matrix for quarterly observations allowing for autocorrelation appears the most plausible. There are no negative elements and the pattern of correlations is intuitively reasonable. We find that Credit Card losses are little correlated with other loss categories except for Other Retail. All Real Estate is fairly highly correlated with other categories except for Credit Cards. The corporate loan categories, Leases, C&I and Agricultural are reasonably highly correlated.

To analyze the pattern of correlation further, we decompose the quarterly correlation matrix, in the two cases with and without autocorrelation, into their principal components (see Table 5). To analyze how far the correlation matrix is from one with a single or two factors, one may examine the magnitudes of the largest eigenvalues. One may also see which of the loan loss categories seem to be driven by the different factors.

The quarterly data results suggest that there are two dominant factors. All the loan loss categories have positive loadings on the first of these, with Leases and C&I having the largest loadings. This factor may be identified with the single credit risk factor presumed by the Basel II parameterization.

The corporate loan categories, C&I and Leases have positive weights on the second factor while the retail categories have negative weights. This second factor may therefore be thought of as differentiating these two broad business sectors.

5.3 Correlation and Capital

As a final exercise, we calculate capital for two portfolios, the first primarily retail and the second primarily corporate. Table 6 shows the portfolio weights for each of the three portfolios on the six different loan categories, (i) a portfolio will equal

weights in the six loan categories, (ii) a predominantly retail loan portfolio, (ii) a predominantly corporate loan portfolio, and (iii) a “combined” portfolio, diversified across these two broad sectors.

Our exercise consists of simulating the model that allows for autocorrelation 1 million times. For each portfolio, we simulate the model under three different assumptions about correlation of factor innovations:

1. We suppose the innovations are perfectly correlated for all loan loss categories. This assumption corresponds to the assumption implicitly used in Basel since under Basel II capital is calculated for each line of business and then added up.
2. We suppose the innovations for different loan loss categories have zero pairwise correlations.
3. We suppose that the correlations of loan loss innovations are those estimated in this paper.

In Table 6, we report VaRs for a confidence level of 99.9% for the four portfolios.

The results show that going from zero to perfect correlation boosts capital by about 70-80%. The estimated degree of correlation gives results that are about half way from zero to perfect correlation. Broadly speaking, these results hold for all of the portfolios although the rise in capital that results from increasing the degree of correlation is greatest, as one might expect in the case of diversified portfolios like our so-called “Combined Portfolio”.

6 Conclusion

This paper presents a generalization of a widely used Vasicek model of loan losses. Our generalization allows for the dynamic evolution of loan loss distributions. We argue that introducing dynamics is essential if such models are to be employed to analyze real life data since loan losses exhibit clear signs of autocorrelation and trends.

Estimating our models using aggregate US loan loss data, we show that the capital charges imposed on banks under Basel II are very conservative in the case of retail exposures. We also examine the impact on economic capital of diversifying across different broad sectors of the credit market. (As is well known, Basel II does not

grant banks concessions in required capital for diversification across broad business lines.)

Our analysis provides a framework and a set of tools for analyzing portfolio credit risk at the aggregate, broad business line level. In future work, we intend to look at conditional modeling of loan losses and its use in stress testing and at modeling correlations for individual loan data.

References

- ALLEN, L., AND A. SAUNDERS (2003): “survey of cyclical effects in credit risk measurement models,” Working Paper 126, Bank for International Settlements.
- ALTMAN, E. I., AND A. SAUNDERS (2001): “An Analysis and Critique of the BIS Proposal on Capital Adequacy Ratings,” *Journal of Banking and Finance*, 25, 25–46.
- BANGIA, A., F. DIEBOLD, A. KRONIMUS, C. SCHAGEN, AND T. SCHUERMAN (2002): “Ratings migration and the business cycle, with application to credit portfolio stress testing,” *Journal of Banking and Finance*, 26(2-3), 445–474.
- BASEL COMMITTEE ON BANKING SUPERVISION (1999): “A New Capital Adequacy Framework,” unpublished mimeo, Bank for International Settlements, Basel.
- BLACK, S., AND D. MORGAN (1998): “Risk and democratisation of credit cards,” Research Paper 8915, Federal Reserve Bank of New York, New York.
- BURTSCHHELL, X., J. GREGORY, AND J.-P. LAURENT (2005): “A Comparative Analysis of CDO Pricing Models,” Working paper, BNP Paribas.
- COUDERC, F., O. RENAULT, AND O. SCAILLET (2003): “Credit risk dynamics: forecasting rating transition matrices,” Working paper, FAMA, Geneva.
- COWAN, A. M., AND C. D. COWAN (2004): “Default correlation: An empirical investigation of a subprime lender,” *Journal of Banking and Finance*, pp. 753–771.
- DAS, S. R., D. DUFFIE, N. KAPADIA, AND L. SAITA (2005): “Common Failings: How Corporate Defaults are Correlated,” mimeo, Stanford University, Stanford.

- DE ANDRADE, F. W. M., AND L. C. THOMAS (2004): “Structural Models in Consumer Credit,” Discussion paper, University of Southampton, UK.
- DIETSCH, M., AND J. PETEY (2004): “Should SME Exposures be Treated as Retail or Corporate Exposures? A comparative Analysis of Default Probabilities and Asset Correlations in the French and German SMEs,” *Journal of Banking and Finance*, 28, 773–788.
- DUFFIE, D., AND N. GARLEANU (2001): “Risk and valuation of collateralized debt obligations,” *Financial Analysts Journal*, 57, 41–59.
- DUNN, L., AND T. KIM (1999): “An Emirical Investigation of Credit Card Default,” Working Paper 44, Ohio State University.
- EMBRECHTS, P., F. LINDSKOG, AND A. MCNEIL (2003): “Modelling Dependence with Copulas and Applications to Risk Management,” in *Handbook of Heavy Tailed Distributions in Finance*, ed. by S. Rachev. Elsevier/North-Holland, Amsterdam.
- FREY, R., AND A. MCNEIL (2003): “Dependent Defaults in Models of Portfolio Credit Risk,” *Journal of Risk*, 6(1), 59–92.
- GORTON, G. B., AND P. HE (2003): “Bank credit cycles,” Working paper, University of Pennsylvania, Wharton School.
- HANSON, S., H. PESARAN, AND T. SCHUERMAN (2006): “Firm Heterogeneity and Credit Risk Diversification,” mimeo, University of Cambridge, Cambridge.
- HEITFIELD, ERIK, S. B., AND S. CHOMSISENGPHET (2005): “The Effects of Name and Sector Concentrations on the Distribution of Losses for Portfolios of Large Wholesale Credit Exposures,” mimeo, Federal Reserve Board, Washington DC.
- HULL, J., AND A. WHITE (2004): “Valuation of a CDO and nth to Default CDS Without Monte Carlo Simulation,” *Journal of Derivatives*, 12(2), 8–23.
- KIESEL, R., AND R. SCHMIDT (2004): “A Survey of Dependency Modelling: Copula, Tail Dependence and Estimation,” in *Structured Credit Products: Pricing, Rating, Risk Management and Basel II*, ed. by W. Perraudin, pp. 3–34. Risk Books, London.
- KOOPMAN, S. J., AND A. LUCAS (2005): “Business and default cycles for credit risk,” *Journal of Applied Econometrics*, 20, 311323.

- KOOPMAN, S. J., A. LUCAS, AND P. KLAASSEN (2005): “Empirical credit cycles and capital buffer formation,” *Journal of Banking and Finance*, forthcoming.
- KUPIEC, P. (2001): “The new Basel Capital Accord: The devil is in the (calibration) details,” Working Paper 113, IMF, Washington D.C.
- LAURENT, J.-P., AND J. GREGORY (2003): “Basket Default Swaps, CDOs and Factor Copulas,” Working paper, ISFA Actuarial School, University of Lyon, Lyon.
- LI, D. X. (2000): “On Default Correlation: A Copula Approach,” *Journal of Fixed Income*, 9, 43–54.
- MCNEIL, A. J. M., AND J. WENDIN (2005a): “Bayesian Inference for Generalized Linear Mixed Models of Portfolio Credit Risk,” Working paper, ETH University of Zurich, Zurich.
- (2005b): “Dependent Credit Migrations,” Working paper, ETH University of Zurich, Zurich.
- NICKELL, P., W. R. PERRAUDIN, AND S. VAROTTO (2000): “The Stability of Ratings Transitions,” *Journal of Banking and Finance*, 24(1-2), 203–227.
- PERLI, R., AND W. I. NEYDA (2004): “Economic and Regulatory Capital Allocation for Revolving Retail Exposures,” *Journal of Banking and Finance*, 28, 789–809.
- SCHONBUCHER, P. J. (2002): “Taken to the Limit: Simple and not so simple Loan Loss Distributions,” Working paper, Bonn University.
- (2004): “Factor Models for Portfolio Credit Risk,” Working paper, Bonn University.
- SCHONBUCHER, P. J. (2005): “Portfolio Losses and the Term Structure of Loss Transition Rates: A New Methodology for the Pricing of Portfolio Credit Derivatives,” mimeo, Eth University of Zurich, Zurich.
- SIDENIUS, J., V. PITERBARG, AND L. ANDERSON (2004): “A New Framework for Dynamic Credit Portfolio Loss Modelling,” mimeo.
- VAROTTO, S. (2005): “The Joys of Industrial Diversification in the Stock and Eurobond Markets,” mimeo, University of Reading, Reading.

VASICEK, O. (1991): “Limiting Loan Loss Probability Distribution,” Unpublished mimeo, KMV.

WEHRSPHON, U. (2003): “Analytic Loss Distributions of Heterogeneous Portfolios in the Asset Value Credit Risk Model,” Working paper, Alfred Weber Institute, Heidelberg University.

WHITLEY, J., R. WINDRAM, AND P. COX (2004): “An Empirical Model of Household Arrears,” Working Paper 214, Bank of England, London.

WILSON, T. (1997a): “Portfolio credit risk, part I,” *Risk*, pp. 111–117.

——— (1997b): “Portfolio credit risk, part II,” *Risk*, pp. 56–61.

Table 1: Parameter Estimates - Quarterly Loan Series

	All Real Estate	Credit Cards Consumer	Other Consumer	Leases	C&I Loans	Agricultural Loans
Series statistics						
Loss rate series mean (%)	0.26	2.27	0.38	0.29	0.50	0.39
Loss rate std dev (%)	0.25	0.34	0.12	0.16	0.31	0.61
Parameter Estimates with No Autocorrelation						
Std dev transformed loss rate (%)	30.72	6.29	10.70	22.44	25.37	38.42
Factor correlation ρ (%)	8.62	0.39	1.13	4.79	6.05	12.86
Std error ρ (%)	1.24	0.06	0.18	0.72	0.89	1.76
Unconditional default probability q (%)	0.26	2.27	0.38	0.29	0.51	0.36
Std error q (%)	0.08	0.07	0.04	0.06	0.11	0.13
Parameter Estimates with Autocorrelation						
Std dev transformed loss rate (%)	12.41	3.11	6.98	14.51	13.56	29.08
AR(1) parameter β (%)	85.67	74.43	52.87	59.37	73.84	40.93
Std error β (%)	8.40	9.75	10.55	11.14	10.35	10.81
Factor correlation ρ (%)	9.70	0.38	1.02	4.93	6.57	12.52
Std error ρ (%)	5.32	0.15	0.28	1.48	2.61	2.65
Unconditional default probability q (%)	0.24	2.28	0.39	0.29	0.48	0.33
Std error q (%)	0.14	0.14	0.03	0.06	0.15	0.09

Table 2: Basel Capital Calculations

	All Real Estate		Credit Cards		Other Consumer		Leases		C&I Loans		Agricultural Loans	
	Estate	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer
Basel Correlation Estimates												
Default probability (%)	0.26	7.30	2.05	0.92	1.18	0.42						
Factor correlation ρ (%)	15.00	4.00	9.34	19.58	18.66	21.72						
Loss given default LGD	0.35	0.65	0.65	0.45	0.45	0.45						
Capital Calculations												
Capital implied with no autocorrelation (%)	1.35	2.40	1.79	2.03	3.12	2.88						
Capital implied with autocorrelation (%)	0.28	2.22	0.90	1.20	1.35	0.70						
Basel implied capital (%)	1.96	8.97	8.07	7.16	7.82	5.14						

Table 3: Normality Tests Applied to Shocks

Normality Test	
Lilliefors Statistic [†]	Jaques-Bera Statistic [‡]
0.136*	4.55
0.093	4.60
0.093	4.93
0.114*	5.10
0.101*	4.92
0.098*	4.74

*Normality hypothesis rejected.

Table 4: Factor Correlation Matrices - Quarterly Loan Series

	All Real Estate	Credit Cards Consumer	Other Consumer	Leases	C&I Loans	Agricultural Loans
Correlation Matrix with No Autocorrelation						
All Real Estate	1.00	0.22	-0.19	0.38	0.58	0.28
Credit Cards Consumer	0.22	1.00	0.42	0.52	0.46	0.22
Other Consumer	-0.19	0.42	1.00	0.55	0.45	0.03
Leases	0.38	0.52	0.55	1.00	0.85	0.53
C&I Loans	0.58	0.46	0.45	0.85	1.00	0.55
Agricultural Loans	0.28	0.22	0.03	0.53	0.55	1.00
Correlation Matrix with Autocorrelation						
All Real Estate	1.00	0.23	0.54	0.57	0.72	0.68
Credit Cards Consumer	0.23	1.00	0.41	0.26	0.24	0.15
Other Consumer	0.54	0.41	1.00	0.47	0.64	0.32
Leases	0.57	0.26	0.47	1.00	0.43	0.46
C&I Loans	0.72	0.24	0.64	0.43	1.00	0.64
Agricultural Loans	0.68	0.15	0.32	0.46	0.64	1.00

Table 5: Principal Component Analysis - Quarterly Loan Series

	Eigenvalue	All Real Estate		Credit Cards		Other Consumer		Leases		C&I Loans		Agricultural Loans	
		Estate	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer
Eigenvalue Decomposition and Common Factor Weights with No Autocorrelation													
Common Factor 1	0.11	1.24	-0.19	0.98	0.91	-2.41	0.57						
Common Factor 2	0.16	-0.69	-0.03	-1.20	1.89	-0.30	-0.72						
Common Factor 3	0.56	-0.30	1.01	-0.44	-0.24	-0.39	0.51						
Common Factor 4	0.74	0.60	0.54	-0.06	-0.13	-0.00	-0.82						
Common Factor 5	1.32	-0.52	0.23	0.59	0.07	-0.09	-0.28						
Common Factor 6	3.11	0.17	0.21	0.17	0.30	0.30	0.20						
Eigenvalue Decomposition and Common Factor Weights with Autocorrelation													
Common Factor 1	0.22	-0.66	0.12	-0.88	0.52	1.61	-0.67						
Common Factor 2	0.27	-1.54	-0.11	0.53	0.20	0.03	1.02						
Common Factor 3	0.60	-0.06	-0.28	0.12	1.10	-0.45	-0.40						
Common Factor 4	0.64	-0.08	-0.52	0.87	-0.30	0.34	-0.58						
Common Factor 5	1.16	0.21	-0.78	-0.25	0.03	0.14	0.36						
Common Factor 6	3.46	0.25	0.15	0.24	0.20	0.24	0.22						

Table 6: Monte Carlo Simulated Portfolio - Quarterly Loan Series

	Portfolio Weights									
	All Real	Credit Cards		Other	Leases	C and I	Agricultural			
	Estate	Consumer	Consumer	Consumer		Loans	Loans	Loans	Loans	Loans
Equal weight portfolio	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Retail portfolio	0.45	0.15	0.15	0.20	0.06	0.10	0.04	0.04	0.04	0.04
Corporate portfolio	0.10	0.03	0.03	0.03	0.40	0.40	0.04	0.04	0.04	0.04
Combined portfolio	0.28	0.09	0.09	0.12	0.23	0.25	0.04	0.04	0.04	0.04
	Perfect Correlation	Zero Correlation	Zero Correlation	Correlated						
	Implied VaR [†]	Implied VaR [†]	Implied VaR [†]	Implied VaR [†]						
Equal weight portfolio	1.73	1.09	1.09	1.35						
Retail portfolio	1.20	0.70	0.70	0.94						
Corporate portfolio	1.37	0.83	0.83	1.06						
Combined portfolio	1.28	0.71	0.71	0.98						

[†]Simulated VaR quantiles at 99.9%.

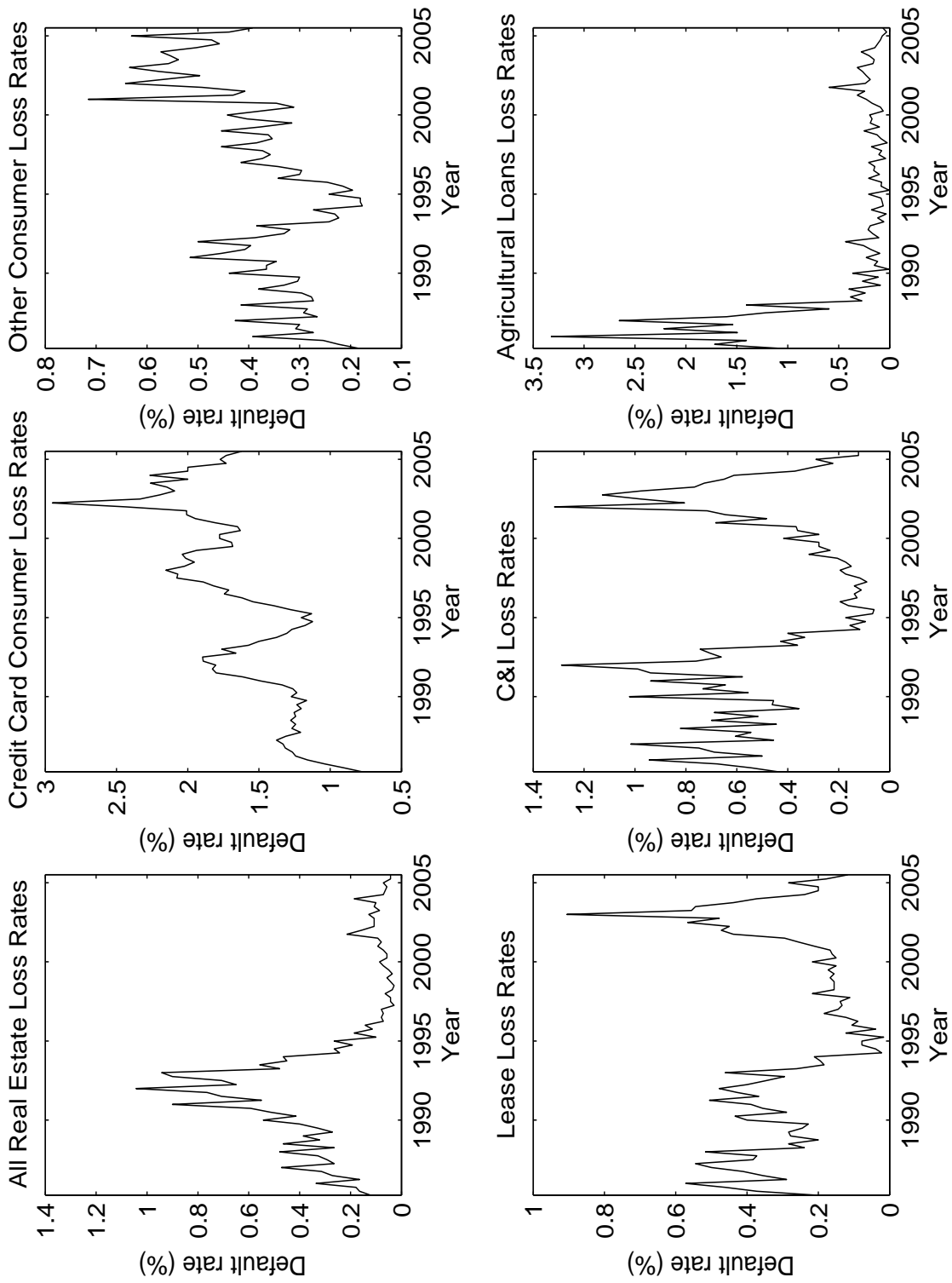


Figure 1: Historical US Bank Quarterly Loan Loss Rates

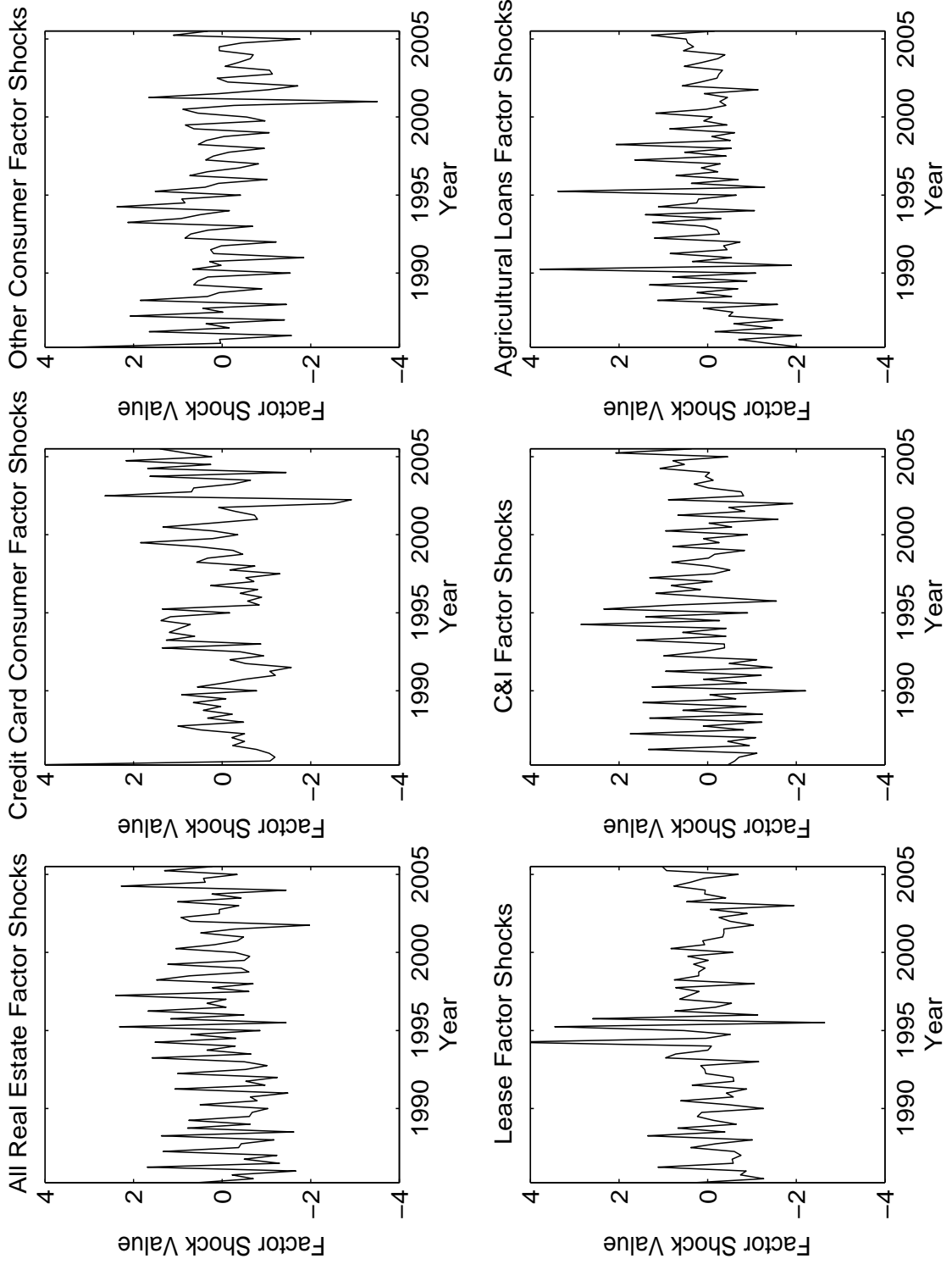


Figure 2: Factor Shocks for US Bank Quarterly Loan Loss Rates