Firm Heterogeneity and Credit Risk Diversification*

Samuel G. Hanson†
Harvard University
shanson@fas.harvard.edu

M. Hashem Pesaran
Faculty of Economics, University of Cambridge
mhp1@cam.ac.uk

Til Schuermann‡
Federal Reserve Bank of New York and Wharton Financial Institutions Center
til.schuermann@ny.frb.org

June 2006

Abstract

This paper examines the impact of *neglected* heterogeneity on credit risk. We show that neglecting heterogeneity in firm returns and/or default thresholds leads to *under*estimation of expected losses (EL), and its effect on portfolio risk is ambiguous. But once EL is controlled for, neglecting parameter heterogeneity leads to *over*estimation of risk. Using a portfolio of U.S. firms we illustrate that heterogeneity in the default threshold or probability of default, measured for instance by a credit rating, is of first order importance in affecting the shape of the loss distribution: including ratings heterogeneity alone results in a 20% drop in loss volatility and a 40% drop in 99.9% VaR, the level to which the risk weights of the New Basel Accord are calibrated.

JEL Classifications: C33, G13, G21.

Key Words: Risk management, correlated defaults, factor models, portfolio choice.

*We would like to thank Richard Cantor, Paul Embrechts, Michael Gordy, Rafael Repullo, Joshua Rosenberg, Jose Scheinkman, Zhenyu Wang, Alan White, and especially David Lando, as well as participants at the 13^{th} Annual Conference on Pacific Basin Finance, Economics, and Accounting at Rutgers University, June 2005, the BIS-ECB 4th Joint Central Bank Research Conference in Frankurt, November 2005, the NBER conference on Risks of Financial Institutions, Cambridge, MA, November 2005, the CIRANO-CIREQ Conference on Financial Econometrics, Montreal, Canada, May 2006, the Financial Econometrics Conference at University of York, UK, June 2006, and seminar participants at the Newton Institute, University of Cambridge, UMass Amherst and the Federal Reserve Bank of New York for helpful comments and suggestions. We are also grateful to Yue Chen and Chris Metli for research assistance.

†PhD Program in Business Economics, Harvard University, Dept. of Economics and Harvard Business School

‡Corresponding author. Any views expressed represent those of the authors only and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Practically all credit risk models to date owe an intellectual debt to the options based approach to firm default by Merton (1974).[1] It took more than a decade for the development of solutions to portfolio loss distributions (Vasicek 1987, 1991), and those solutions were obtained under strict homogeneity assumptions regarding the probability distribution of firms asset values and default thresholds. Yet clearly heterogeneity is important and has received considerable attention of late. Two notable examples are Gordy (2003) using a factor-based approach,[2] and Duffie, Saita and Wang (2005) using a default intensity approach.

To take full account of firm heterogeneity in credit risk places great demands on the data. When firms are public and have traded securities such as stocks, bonds, or even credit default swaps (CDS), as well as third party assessments such as credit ratings, there is great scope for allowing and accounting for heterogeneity. But this scenario is limited to a small minority of firms; indeed most loans in banks' portfolios are to privately held firms about which we (and the banks) know rather little. In that case one may be forced to settle for the credit portfolio solutions obtained under homogeneity. What then are the consequences of *neglecting* heterogeneity for the analysis of the loss distribution? What is the impact on expected loss (EL), on risk, whether measured by loss volatility, which we call unexpected loss (UL), or tail quantiles (value at risk, VaR), or the shape of the entire loss distribution? Moreover, which sources of heterogeneity are especially important?

This is the focus of our paper, and to our knowledge we are the first to examine the impact of neglected heterogeneity on credit risk. We consider both observed and unobserved types of heterogeneity. The former is relatively easy to deal with and does not pose any particular technical difficulties.[3] The latter (unobserved heterogeneity) is more difficult and will be the focus of our analysis. Note that parameter heterogeneity refers to differences in population values of the parameters across different firms and prevails even in the absence of estimation uncertainty.[4] We build on the work of Vasicek and Gordy and examine the consequences of incorrectly neglecting the heterogeneity of return correlations and default thresholds across firms for the analysis of loss distributions. The default threshold captures a variety of firm characteristics such as balance sheet structure, including leverage, as well as intangibles like the quality of management. This heterogeneity can be random – firms, say, have on average the same factor loadings – and/or the differences could be systematic – mean factor loadings could differ across industries but are randomly distributed around the industry mean, across firms within an industry.

[1] For a summary of models see Saunders and Allen (2002), and for detailed comparisons, see Koyluoglu and Hickman (1998) and Gordy (2000).

[2] Gordy's (2003) result shaped regulatory policy in the specific form of the regulatory capital formula in the New Basel Accord (BCBS 2005, §272).

[3] In our set up an important example of observed heterogeneity is given by credit ratings across firms.

[4] In this paper we do not allow for parameter estimation uncertainty.

Our theoretical set-up is quite general and imposes few distributional and parametric restrictions. The theoretical results show a complex interaction between the sources of heterogeneity and the resulting loss distribution. We find that incorrectly neglecting heterogeneity results in *under*estimation of expected losses, and its effect on portfolio risk is ambiguous. This is a new result and arises due to the nonlinear nature of the relationships that prevail between the return process, the default threshold and the resultant default (and hence loss) process. Differences in asset values and default thresholds across firms do not disappear by cross-section averaging even if the differences across firms are random and the underlying portfolio is sufficiently large.

In comparing heterogeneous loss portfolios it is therefore important that appropriate adjustments are made so that the different portfolios all have the same EL's. This is only possible by allowing for systematic heterogeneity across firms, e.g. by grouping firm into industries, regions, distances to default (e.g. credit rating), or a combination of those. In that case we prove that neglected heterogeneity results in *over*estimated risk, so that falsely imposing homogeneity can be quite costly.

Along the way we derive analytic solutions to loss distributions under parameter heterogeneity, assuming that the cross-section means and variance/covariances of the firm parameters are known; under homogeneity these variances and covariances are set to zero. Such derivations are important since they allow us to calibrate loss distributions for cases where there is little or no data to estimate the extent of parameter heterogeneity (which is practically the case for most of bank lending), using available estimates based on publicly traded securities. The latter estimates are not perfect and will be subject to errors, but are likely to be more appropriate than setting the variance and covariances to zero. This result marks our second contribution to the literature.

The importance of these theoretical insights are illustrated using a portfolio of about 600 publicly traded U.S. firms. Return regressions subject to different degrees of parameter heterogeneity are estimated recursively using six ten-year rolling estimation windows, and for each estimation window the loss distribution is then simulated out-of-sample over a one-year period. The predictions made by theory are confirmed in this application and are found to be robust across the six years. We show that heterogeneity in the default threshold or probability of default (PD), measured for instance by a credit rating, is of first order importance in affecting the shape of the loss distribution: allowing for ratings heterogeneity alone results in a 20% drop in loss volatility (keeping EL's constant) and 40% drop in 99.9% VaR, the level to which the risk weights in the New Basel Accord are calibrated. Allowing for additional heterogeneity results in a further 10% drop in 99.9% VaR. This result has important policy implications as a PD estimate through a credit rating, whether generated by a bank internally or provided by a rating agency externally, is the one parameter (of those considered here) that is allowed to vary in the New Basel Accord.

To analyze the impact of neglected heterogeneity on credit risk, we use a simple multifactor approach which is easily adapted to this task. Multifactor models have been used extensively in finance following Ross (1976) and Chamberlain and Rothschild (1983).[5] Their application to credit risk has been more recent. A notable example is its use in the CreditMetrics model as set out in Gupton, Finger and Bhatia (1997). Gordy (2000) and Schönbucher (2003, ch. 10) provide useful reviews.

A separate line of research has focused on correlated default intensities as in Lando (1998), Schönbucher (1998), Duffie and Singleton (1999), Duffie and Gârleanu (2001), Collin-Dufresne, Goldstein and Hugonnier (2004), and Duffie, Saita and Wang (2005); with a review by Duffie (2005). There are also a host of other approaches, including correlated (but non-systematic) jumps-at-default (Driessen 2005, Jarrow, Lando, and Yu 2005), the contagion model of Davis and Lo (2001) as well as Giesecke and Weber's (2004) indirect dependence approach, where default correlation is introduced through local interaction of firms with their business partners as well as via global dependence on economic risk factors. The idea of generalizing default dependence beyond correlation using copulas is discussed in Li (2000), Embrechts, McNeil, and Straumann (2001), Schönbucher (2002), Frey and McNeil (2003), and Hull and White (2006).

In short, the literature on modeling default dependence is growing rapidly along different paths, and there is as yet no consensus which approach is best. Our paper does not address that issue, but it does highlight, using a factor approach, the impact of neglected heterogeneity. This issue of neglected heterogeneity clearly also arises in the case of other approaches that focus on correlated default intensities or copulas; we leave that for others to explore. The factor structure considered here does allow us to explore two distinct channels of heterogeneity: one that is shared, namely factor sensitivities, and one which is specific to firms within a given grouping (e.g. credit rating), namely the default threshold or the distance to default.

Our results have bearing on risk and capital management as well as the pricing of credit assets. For example, in the case of a commercial bank, ignoring heterogeneity may result in underprovisioning for loan losses since EL is underestimated, and may result in overcapitalization against (bank) default since risk is overestimated. The risk assessment and pricing of complex credit asset such as collateralized debt obligations (CDOs) may be adversely affected since they are driven by the shape of the loss distribution which is then segmented into tranches.

The plan for the remainder of the paper is as follows: Section 2 introduces the basic model of firm value and default and considers the problem of correlated defaults. Section 3 derives the portfolio loss distribution under different heterogeneity assumptions, starting with the simple case of a homogeneous portfolio as introduced by Vasicek. These results are illustrated in Section 4 where we explore the impact of heterogeneity using returns for a large sample of publicly traded firms in the U.S. across seven sectors, and we analyze the resulting loss distributions obtained by stochastic simulations. Section 5 provides some concluding remarks. A technical Appendix presents generalizations of some material in Sections 2 and 3.

[5] Connor and Korajczyk (1995) provide an excellent survey.

2 Firm Value, Default and Default Dependence

Much of the research on credit risk modeling, including our own, is based on the option theoretic default model of Merton (1974). Merton recognized that a lender is effectively writing a put option on the assets of the borrowing firm; owners and owner-managers (i.e. shareholders) hold the call option. If the value of the firm falls below a certain threshold, the owners will put the firm to the debt-holders. Thus a firm is expected to default when the value of its assets falls below a threshold value determined by its liabilities.[6]

2.1 Firm Value and Default

Consider a firm i having asset value V_{it} at time t, and an outstanding stock of debt, D_{it}. Under the Merton model default occurs at the maturity date of the debt, $t+h$, if the firm's assets, $V_{i,t+h}$, are less than the face value of the debt at that time, $D_{i,t+h}$. The value of the firm at time t is the sum of debt and equity, namely

$$V_{it} = D_{it} + E_{it}, \text{ with } D_{it} > 0. \tag{1}$$

Conditional on time t information, default will take place at time $t+h$ if $V_{i,t+h} \leq D_{i,t+h}$. In the Merton model debt is assumed to be fixed over the horizon h. For simplicity we set $h = 1$; extensions to multiple periods can be found in Pesaran, Schuermann, Treutler, and Weiner (2005), hereafter PSTW. Because default is costly and violations to the absolute priority rule in bankruptcy proceedings are common, in practice debtholders have an incentive to put the firm into receivership even before the equity value of the firm hits the zero value.[7] Similarly, the bank might also have an incentive of forcing the firm to default once the firm's equity falls below a non-zero threshold.[8] Importantly, default in a credit relationship is typically a weaker condition than outright bankruptcy. An obligor may meet the technical default condition, e.g. a missed coupon payment, without subsequently going into bankruptcy. As a result we shall assume that default takes place if

$$0 < E_{i,t+1} < C_{i,t+1}, \tag{2}$$

where $C_{i,t+1}$ is a positive default threshold which could vary over time and with the firm's characteristics (such as region or industry sector). Natural candidates that affect the default threshold

[6] An alternative to Merton's end of period approach are the first-passage models where default would occur the first time that firm value falls below a default boundary (or threshold) over the period, as in Zhou (2001).

[7] See, for instance, Leland and Toft (1996) who develop a model where default is determined endogenously, rather than by the imposition of a positive net worth condition. More recently, Broadie, Chernov, and Sundaresan (2005) show that in the presence of APR default can be optimal when $E_{it} > 0$ even in the case of a single debt class.

[8] For a treatment of this scenario, see Garbade (2001).

include observable factors such as leverage, profitability, and firm age (most of which appear in models of firm default), as well as non-observable ones such as management quality.[9]

We are now in a position to consider the evolution of firm equity value which we assume follows a standard geometric random walk model:

$$\ln(E_{i,t+1}) = \ln(E_{it}) + \mu_i + \xi_{i,t+1}, \qquad \xi_{i,t+1} \sim iidN(0, \sigma^2_{\xi_i}), \tag{3}$$

with a non-zero drift, μ_i, and idiosyncratic Gaussian innovations with a zero mean and firm-specific volatility, σ_{ξ_i}. Consequently, default occurs if

$$\ln(E_{i,t+1}) = \ln(E_{i,t}) + \mu_i + \xi_{i,t+1} < \ln\left(C_{i,t+1}\right), \tag{4}$$

or if the one-period change in equity value or return falls below some threshold defined by

$$\ln\left(\frac{E_{i,t+1}}{E_{it}}\right) < \ln\left(\frac{C_{i,t+1}}{E_{it}}\right) = \lambda_{i,t+1}. \tag{5}$$

Equation (5) tells us that the relative (rather than absolute) decline in firm value must be large enough over the period to result in default. Note that firm-specific information such as leverage and management quality, embedded in the default threshold C_i, carry over to λ_i. Thus for highly levered firms with poor management, the threshold is lower (in the sense of being more negative) than for well capitalized and well managed firms. The important issue of measuring λ_i empirically is taken up in Section 4.1.

Under the assumption of Gaussian innovations in (3), the probability that firm i defaults at the end of the period is given by

$$\pi_{i,t+1} = \Phi\left(\frac{\lambda_{i,t+1} - \mu_i}{\sigma_{\xi_i}}\right), \tag{6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. In the theoretical discussions that follows we shall assume that the firm-specific default thresholds are given.

2.2 Cross Firm Default Dependence: Some Preliminaries

In the context of the Merton model, cross firm default dependence can be introduced by assuming that shocks to the value of a firm's equity, $\xi_{i,t+1}$, defined by (3), have the following multifactor structure:[10]

$$\xi_{i,t+1} = \boldsymbol{\gamma}_i' \mathbf{f}_{t+1} + \sigma_i \varepsilon_{i,t+1}, \quad \varepsilon_{i,t+1} \sim iid(0, 1) \tag{7}$$

where $\mathbf{f}_{t+1}$ is an $m \times 1$ vector of common factors, $\boldsymbol{\gamma}_i$ is the associated vector of factor loadings, and $\varepsilon_{i,t+1}$ is the firm-specific idiosyncratic shock, assumed to be distributed independently across i; in

[9] For models of bankruptcy and default at the firm level, see, for instance, Altman (1968), Lennox (1999), Shumway (2001), and Hillegeist, Keating, Cram and Lundstedt (2004).

[10] We consider a simple linear model, though nonlinear factor models with the possibility of jumps are also possible.

this way the model in (7) is said to be conditionally independent.[11] The common factors could be treated as either unobserved or observed through macroeconomic variables such as output growth, inflation, interest rates and exchange rates.[12]

In what follows we suppose the factors are unobserved, distributed independently of $\varepsilon_{i,t+1}$, and have constant variances.[13] Thus, without loss of generality we assume that $\mathbf{f}_{t+1} \sim (\mathbf{0}, \mathbf{I}_m)$, where $\mathbf{I}_m$ is an identity matrix of order m.

Using (7) in (3) we now have

$$\ln(E_{i,t+1}) - \ln(E_{it}) = r_{i,t+1} = \mu_i + \boldsymbol{\gamma}_i' \mathbf{f}_{t+1} + \sigma_i \varepsilon_{i,t+1}. \tag{8}$$

Under the above assumptions

$$\sigma_{\xi_i}^2 = \boldsymbol{\gamma}_i' \boldsymbol{\gamma}_i + \sigma_i^2, \tag{9}$$

which decomposes the return variance into the part due the systematic risk factors, $\boldsymbol{\gamma}_i' \boldsymbol{\gamma}_i$, and the residual or idiosyncratic variance, σ_i^2. The presence of the common factors also introduces a varying degree of asset return correlations across firms, which in turn leads to variation in cross firm default correlations for a given set of default thresholds, $\lambda_{i,t+1}$. The extent of default correlation depends on the size of the factor loadings, $\boldsymbol{\gamma}_i$, the importance of the idiosyncratic shocks, σ_i, the values of the default thresholds, $\lambda_{i,t+1}$, and the shape of the distribution assumed for $\varepsilon_{i,t+1}$, particularly its left tail properties. The correlation coefficient of returns of firms i and j is given by

$$\rho_{ij} = \frac{\boldsymbol{\delta}_i' \boldsymbol{\delta}_j}{\left(1 + \boldsymbol{\delta}_i' \boldsymbol{\delta}_i\right)^{1/2} \left(1 + \boldsymbol{\delta}_j' \boldsymbol{\delta}_j\right)^{1/2}}, \tag{10}$$

where $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i / \sigma_i$ is the standardized $m \times 1$ vector of factor loadings (systematic risk exposures) of firm i.

To derive the cross correlation of firm defaults, which we denote by $\rho_{ij,t+1}^*$, let $z_{i,t+1}$ be the default outcome for firm i over a single period such that

$$z_{i,t+1} = I\left(\lambda_{i,t+1} - r_{i,t+1}\right), \tag{11}$$

where $I(A)$ is an indicator function that takes the value of unity if $A \geq 0$, and zero otherwise. Then (see also Zhou 2001)

$$\rho_{ij,t+1}^* = \frac{E\left(z_{i,t+1} z_{j,t+1}\right) - \pi_{i,t+1}\pi_{j,t+1}}{\sqrt{\pi_{i,t+1}(1 - \pi_{i,t+1})}\sqrt{\pi_{j,t+1}(1 - \pi_{j,t+1})}}, \tag{12}$$

[11] Note that conditional independence may not necessarily be attained in an empirical setting (see, for instance, Das, Duffie, Kapadia and Saita 2005), a point we discuss in more detail in Section 4.

[12] For instance, PSTW provide an empirical implementation of this model by linking the (observable) factors, $\mathbf{f}_{t+1}$, to the variables in a global vector autoregressive model comprising around 80% of world output.

[13] The more general case where the factors may exhibit time varying volatility can be readily dealt with by allowing the factor loadings to vary over time, in line with market volatilities. But in this paper we shall not pursue this line of research, primarily because the focus of our empirical analysis is on quarterly and annual default risks, and over such horizons asset return volatility dynamics tend to be rather weak and of second order importance.

where $\pi_{i,t+1} = E\left(z_{i,t+1}\right)$ is firm $i's$ default probability over the period t to $t+1$. For given values of the thresholds, $\lambda_{i,t+1}$, a relatively simple expression for $\rho^*_{ij,t+1}$ can be obtained if conditional on $\mathbf{f}_{t+1}$, $\varepsilon_{i,t+1}$ and $\varepsilon_{j,t+1}$ are cross sectionally independent, and $\mathbf{f}_{t+1}$ and $\varepsilon_{i,t+1}$ have a joint Gaussian distribution. In this case, known as *conditionally independent double-Gaussian* model, we have

$$\pi_{i,t+1} = \Phi\left(\frac{\lambda_{i,t+1} - \mu_i}{\sqrt{\sigma_i^2 + \boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i}}\right). \tag{13}$$

The argument of $\Phi(\cdot)$ in (13) is commonly referred to as "distance to default" (DD) such that

$$DD_{i,t+1} = \Phi^{-1}(\pi_{i,t+1}) = \frac{\lambda_{i,t+1} - \mu_i}{\sqrt{\sigma_i^2 + \boldsymbol{\gamma}_i'\boldsymbol{\gamma}_i}}. \tag{14}$$

For future reference note that under the double-Gaussian assumption $E\left(z_{i,t+1}z_{j,t+1}\right)$ is given by

$$\begin{aligned} E\left(z_{i,t+1}z_{j,t+1}\right) &= E\left[I\left(\lambda_{i,t+1} - r_{i,t+1}\right)I\left(\lambda_{i,t+1} - r_{i,t+1}\right)\right] \\ &= \Pr\left[r_{i,t+1} < \lambda_{i,t+1} \;\&\; r_{j,t+1} < \lambda_{j,t+1}\right] \\ &= \Phi_2\left[\Phi^{-1}\left(\pi_{i,t+1}\right), \Phi^{-1}\left(\pi_{j,t+1}\right), \rho_{ij}\right], \end{aligned} \tag{15}$$

where $\Phi_2[\cdot]$ is the bivariate standard normal cumulative distribution function, so that the corresponding default correlation ((15) in (12)) is

$$\rho^*\left(\pi, \rho\right) = \frac{\Phi_2\left[\Phi^{-1}\left(\pi\right), \Phi^{-1}\left(\pi\right), \rho\right] - \pi^2}{\pi(1-\pi)}.$$

3 Losses in a Credit Portfolio

Consider now a credit portfolio composed of N different credit assets such as loans, each with exposures or weights w_{it}, at time t, for $i = 1, 2, .., N$, such that[14]

$$\sum_{i=1}^{N} w_{it} = 1,\ \sum_{i=1}^{N} w_{it}^2 = O\left(N^{-1}\right),\ w_{it} \geq 0. \tag{16}$$

A sufficient condition for (16) to hold is given by $w_{it} = O\left(N^{-1}\right)$, which is the standard granularity condition where no single exposure dominates the portfolio.[15] Without loss of generality, we impose both here and later in the empirical section, that a defaulted asset has no recovery value.[16] Under

[14] The assumption that N is time-invariant is made for simplicity and can be relaxed.

[15] Conditions (16) on the portfolio weights was in fact embodied in the initial proposal of the New Basel Accord in the form of the Granularity Adjustments which was designed to mitigate the effects of significant single-borrower concentrations on the credit loss distribution (BCBS, 2001, Ch.8). See also the discussion in Lucas, Klaassen, Sprei, and Straetmans (2001) and Gordy (2004).

[16] The case where default and recovery are correlated through common business cycle effects presents new technical difficulties and is addressed briefly in Appendix A of an earlier version of this paper, available at http://fic.wharton.upenn.edu/fic/papers/05/p0505.html.

this set-up the portfolio loss over the period t to $t+1$ is given by

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it} z_{i,t+1}. \tag{17}$$

The probability distribution function of $\ell_{N,t+1}$ can now be derived both conditional on an information set available at time t, $\mathcal{I}_t$, or unconditionally. The two types of distributions coincide when the factors, $\mathbf{f}_{t+1}$, are assumed to be serially independent, a case often maintained in the literature. However, this assumption precludes the use of any business cycle models in the analysis of credit risk. For the theoretical results we therefore consider the more general case of a dynamic factor model and allow the factors to be serially correlated. In particular, we shall assume that $\mathbf{f}_{t+1}$ follows a covariance stationary process, and $\mathcal{I}_t$ contains at least $\mathbf{f}_t$ and its lagged values, or their determinants when they are unobserved. This structure corresponds to the empirical application in PSTW which makes use of a global macroeconometric model, though later in this paper (Section 4) we impose serial independence on the factor process for expositional simplicity.

A simple example of a dynamic factor model is the Gaussian vector autoregressive specification

$$\mathbf{f}_{t+1} = \mathbf{\Lambda f}_t + \boldsymbol{\eta}_{t+1},\ \ \boldsymbol{\eta}_{t+1} \mid \mathcal{I}_t \sim iidN(\mathbf{0}, \mathbf{\Omega}_{\eta\eta}), \tag{18}$$

where $\mathcal{I}_t$ is the public information known at time t, and $\mathbf{\Lambda}$ is an $m \times m$ matrix of fixed coefficients with all its eigenvalues inside the unit circle such that

$$Var\left(\mathbf{f}_{t+1} \mid \mathcal{I}_t\right) = \sum_{s=0}^{\infty} \mathbf{\Lambda}^s \mathbf{\Omega}_{\eta\eta} \mathbf{\Lambda}'^s = \mathbf{I}_m. \tag{19}$$

The focus of our analysis will be on the limit distribution of $\ell_{N,t+1} \mid \mathcal{I}_t$, as $N \to \infty$. Not surprisingly, this limit distribution depends on the nature of the return process $\{r_{i,t+1}\}$ and the extent to which the returns are cross-sectionally correlated. Our theoretical discussion shall be in terms of the variance of the loss distribution, though occasionally we refer to the standard deviation or loss volatility, known as unexpected loss (UL) in the credit risk literature. In practice, one may also be interested in quantiles of the loss distributions, or VaR, and those can be easily obtained through stochastic simulations.

3.1 Credit Risk under Firm Homogeneity

Vasicek (1987) was one the first to consider the limit distribution of $\ell_{N,t+1}$ using asset return equations with a factor structure. However, he focused on the perfectly homogeneous case with the same factor loadings, $\boldsymbol{\gamma}_i = \boldsymbol{\gamma}$, the same default thresholds, $\lambda_{i,t+1} = \lambda$, the same firm-specific volatilities, $\sigma_i = \sigma$, and zero unconditional returns, $\mu_i = 0$, for all i and t. Note that a multifactor model with homogeneous factor loadings is equivalent to a single factor model. In this model the pair-wise asset return correlations, ρ_{ij}, is identical for all obligor pairs in the portfolio, so that

$$r_{i,t+1} = \sqrt{\rho} f_{t+1} + \sqrt{1-\rho}\,\varepsilon_{i,t+1}, \qquad \begin{pmatrix} \varepsilon_{i,t+1} \\ f_{t+1} \end{pmatrix} \mid \mathcal{I}_t \sim iidN\left(\mathbf{0}, \mathbf{I}_2\right). \tag{20}$$

The remaining parameter, λ, is then calibrated to a pre-specified default probability, π, so that the distance to default and default thresholds are the same for all firms and can be easily estimated from historical default frequency of the portfolio using

$$\lambda = DD = \Phi^{-1}\left(\pi\right). \tag{21}$$

When default thresholds are allowed to vary across firms, identification issues arise which are discussed in Section 4.1.

Under the Vasicek model portfolio loss variance depends on π and ρ^*:

$$Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \pi(1-\pi)\left\{\rho^* + (1-\rho^*)\sum_{j=1}^{N} w_{jt}^2\right\}. \tag{22}$$

Under the granularity condition, (16), for N sufficiently large the second term in brackets becomes negligible. Hence, in the limit

$$\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \pi(1-\pi)\rho^* = \Phi_2\left[\Phi^{-1}\left(\pi\right), \Phi^{-1}\left(\pi\right), \rho\right] - \pi^2. \tag{23}$$

Vasicek's credit loss limit distribution is fully determined by two parameters, namely the average default probability, π, and the pair-wise return correlation coefficient, ρ (see Appendix A.2 for further detail). The former fixes the expected loss of the portfolio, while the latter controls the shape of the loss distribution. In effect one parameter, ρ, controls all aspects relating to the shape of the loss distribution: its volatility, skewness and kurtosis.

3.2 Credit Risk with Firm Heterogeneity

Building on Vasicek's work we now consider models that allow for firm heterogeneity across a number of relevant parameters. In this section we provide some analytical derivations and show how the theoretical work of Vasicek can be generalized. An empirical evaluation of the importance of allowing for firm heterogeneity in credit risk analysis is discussed in Section 4.

Under the heterogeneous multifactor return process (8), the portfolio loss, $\ell_{N,t+1}$, can be written as

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it} I\left(a_{i,t+1} - \boldsymbol{\delta}_i' \mathbf{f}_{t+1} - \varepsilon_{i,t+1}\right), \tag{24}$$

where

$$a_{i,t+1} = \frac{\lambda_{i,t+1} - \mu_i}{\sigma_i}, \quad \boldsymbol{\delta}_i = \frac{\boldsymbol{\gamma}_i}{\sigma_i}. \tag{25}$$

In addition to allowing for parameter heterogeneity, we also relax the assumption that conditional on $\mathcal{I}_t$ the common factors, $\mathbf{f}_{t+1}$, and the idiosyncratic shocks, $\varepsilon_{i,t+1}$, are normally distributed with zero means. Accordingly we assume that

$$\begin{aligned} \varepsilon_{i,t+1} \quad &| \quad \mathcal{I}_t \sim iid\ (0,1),\ \text{ for all } i \text{ and } t, \\ \mathbf{f}_{t+1} \quad &| \quad \mathcal{I}_t \sim iid\ (\boldsymbol{\mu}_{ft}, \mathbf{I}_m),\ \text{ for all } t, \end{aligned}$$

where under the dynamic factor model (18), $\boldsymbol{\mu}_{ft} = \Lambda \mathbf{f}_t$. Allowing $\boldsymbol{\mu}_{ft}$ to be time-varying enables us to explicitly consider the possible effects of business cycle variations on the loss distribution. In the credit risk literature $\boldsymbol{\mu}_{ft}$ is usually set to zero. For future use we shall denote the $\mathcal{I}_t$-conditional probability density and the cumulative distribution functions of $\varepsilon_{i,t+1}$ and $\mathbf{f}_{t+1}$, by $f_\varepsilon(\cdot)$ and $F_\varepsilon(\cdot)$, and $f_{\mathbf{f}}(\cdot)$ and $F_{\mathbf{f}}(\cdot)$, respectively.

To deal with parameter heterogeneity across firms we abstract from time variations in the default thresholds (namely set $a_{i,t+1} = a_i$) and adopt the following random coefficient model:

$$\boldsymbol{\theta}_i = \boldsymbol{\theta} + \mathbf{v}_i,\ \ \mathbf{v}_i \backsim iid\ (\mathbf{0}, \Omega_{vv}),\ \text{ for } i = 1, 2, ..., N, \tag{26}$$

where

$$\boldsymbol{\theta}_i = \left(a_i, \boldsymbol{\delta}_i'\right)',\ \boldsymbol{\theta} = \left(a, \boldsymbol{\delta}'\right)',\ \mathbf{v}_i = \left(v_{ia}, \mathbf{v}_{i\delta}'\right)', \tag{27}$$

and

$$\boldsymbol{\Omega}_{vv} = \begin{pmatrix} \omega_{aa} & \boldsymbol{\omega}_{a\boldsymbol{\delta}} \\ \boldsymbol{\omega}_{\boldsymbol{\delta}a} & \boldsymbol{\Omega}_{\boldsymbol{\delta\delta}} \end{pmatrix}, \tag{28}$$

is a positive semi-definite symmetric matrix, and $\mathbf{v}_i$'s are distributed independently of $(\varepsilon_{j,t+1},\ \mathbf{f}_{t+1})$ for all i, j and t.

Allowing for such parameter heterogeneity may be desirable when firms have different sensitivities to the systematic risk factors $\mathbf{f}_{t+1}$, and those sensitivities or factor loadings are known only up to their distributional properties described in (26). A practical example might be assessing the credit risk for a portfolio of borrowers which are privately held, i.e. not publicly traded. This is typically the case for much of middle market and most of small business lending. For such firms it would be very difficult or even impossible to obtain individual estimates of $\boldsymbol{\theta}_i$, and an average estimate based on $\boldsymbol{\theta}$ and $\boldsymbol{\Omega}_{vv}$ may need to be used. See also Section 4.6.

The heterogeneity described in (26) to (28) states that firm differences are purely random. However, firms could in addition exhibit systematic parameter differences, say by industry and/or region, so that parameter means and covariances are also industry and/or region specific. This generalization is taken up in Section 3.4.2.

3.3 Limits to Unexpected Loss under Parameter Heterogeneity

The extent to which credit losses are diversifiable can be investigated using a number of different measures. Before exploring the entire loss distribution, for reasons of analytical tractability we focus here on loss variance, $Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$, or its square root, unexpected loss, and note that in general

$$Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = E_f\left[Var\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right] + Var_f\left[E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right]. \tag{29}$$

Because of the dependence of the default indicators, $z_{i,t+1}$, across i, through the common factors $\mathbf{f}_{t+1}$, unexpected loss remains even for a portfolio of infinitely many exposures. The problem of correlated defaults can be dealt with by first conditioning the analysis on the source of cross-dependence (namely $\mathbf{f}_{t+1}$) and noting that conditional on $\mathbf{f}_{t+1}$ the default indicators, $z_{i,t+1} = I\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1} - \varepsilon_{i,t+1}\right)$, $i = 1, 2, ..., N$, are independently distributed. Since the $z_{i,t+1}$ are conditionally independent, under granularity condition (16), $E\left[Var\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right] \to 0$, as $N \to \infty$, and in the limit the loss variance, $Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$, is dominated by the second term in (29). Namely, we have

$$\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \lim_{N\to\infty}\left\{Var\left[E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right]\right\}, \tag{30}$$

which follows from Proposition 2 in Gordy (2003). This result clearly shows that when the portfolio weights satisfy the granularity condition (16), the limit behavior of the unexpected loss does not depend on the portfolio weights w_{it}. Furthermore, this result holds irrespective of whether a_i and $\boldsymbol{\delta}_i$ are homogeneous or vary randomly across i.

Under the random coefficient model (26), asymptotic loss variance, given by (30), can be obtained by integrating out the heterogeneous effects of a_i and $\boldsymbol{\delta}_i$. First note that $\ell_{N,t+1} = \sum_{i=1}^N w_{it} I\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1} - \varepsilon_{i,t+1}\right)$, which under (26) can be written as

$$\ell_{N,t+1} = \sum_{i=1}^N w_{it} I\left(a - \boldsymbol{\delta}'\mathbf{f}_{t+1} - \zeta_{i,t+1}\right), \tag{31}$$

where

$$\zeta_{i,t+1} = \varepsilon_{i,t+1} - \mathbf{v}_i'\mathbf{g}_{t+1} \tag{32}$$

captures all innovations, and $\mathbf{g}_{t+1} = (1, -\mathbf{f}_{t+1}')'$. Conditional on $\mathbf{f}_{t+1}$, $\zeta_{i,t+1}$ is distributed independently across i with zero mean and variance

$$\omega_{t+1}^2 = 1 + \mathbf{g}_{t+1}'\boldsymbol{\Omega}_{vv}\mathbf{g}_{t+1}, \tag{33}$$

where $\mathbf{g}_{t+1}'\boldsymbol{\Omega}_{vv}\mathbf{g}_{t+1}$ is the variance contribution arising from the random coefficients model (i.e. explicitly due to parameter heterogeneity). The expected loss conditional on $\mathbf{f}_{t+1}$ is given by

$$\begin{aligned} E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) &= \sum_{i=1}^N w_{it} \Pr\left(\zeta_{i,t+1} \le a - \boldsymbol{\delta}'\mathbf{f}_{t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) \\ &= \sum_{i=1}^N w_{it} F_{\varkappa}\left(\frac{\boldsymbol{\theta}'\mathbf{g}_{t+1}}{\omega_{t+1}}\right), \end{aligned}$$

and since $\sum_{i=1}^{N} w_{it} = 1$, then

$$E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = F_{\varkappa}\left(\frac{\boldsymbol{\theta}'\mathbf{g}_{t+1}}{\omega_{t+1}}\right), \tag{34}$$

where $F_{\varkappa}(\cdot)$ is the cumulative distribution function of the standardized composite innovations

$$\varkappa_{i,t+1} = \frac{\zeta_{i,t+1}}{\omega_{t+1}} \mid \mathbf{f}_{t+1}, \mathcal{I}_t \sim iid(0,1). \tag{35}$$

The loss distribution (34) describes the general case of parameter heterogeneity, and evaluation such as computing EL and VaR, may be done using stochastic simulation by taking independent draws from any given distribution of $\varkappa_{i,t+1}$. In some cases we are able to make predictions analytically, e.g. when heterogeneity is limited to mean returns and/or default thresholds, or to the factor loadings. Those cases are taken up in Section 3.4.

In the limit, therefore, using (30) we have

$$\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = Var\left[F_{\varkappa}\left(\frac{\boldsymbol{\theta}'\mathbf{g}_{t+1}}{\omega_{t+1}}\right) \mid \mathcal{I}_t\right], \tag{36}$$

which does not depend on the exposure weights, w_{it}. This result represents a generalization of the limit variance obtained for the homogeneous case, given above by (23).

The implication for credit risk management is clear: changing the exposure weights that satisfy the granularity condition (16) will have no risk diversification impact so long as all firms in the portfolio have the same risk factor loading distribution. To achieve systematic diversification one needs different firm types, e.g. along industry lines, and we treat this case below in Section 3.4.1.

3.4 Impact of Neglected Heterogeneity

Parameter heterogeneity can significantly affect the shape of the loss distribution as well as expected and unexpected losses. This is most easily illustrated with a single factor model. Multifactor generalizations are given in Appendix A. As before, portfolio losses are given by (replacing w_{it} with w_i to simplify the notation)

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_i I\left(a - \delta f_{t+1} - \zeta_{i,t+1}\right),$$

where $a = (\lambda - \mu)/\sigma$, $\delta = \gamma/\sigma$, and $\zeta_{i,t+1} = \varepsilon_{i,t+1} - v_{ia} + v_{i\delta} f_{t+1}$, is the composite innovation. In the absence of heterogeneity, δ and a can be written in terms of the return correlation, ρ, and the default probability, π :

$$\delta = \sqrt{\frac{\rho}{1-\rho}}, \text{ for } \rho > 0, \tag{37}$$

and

$$a = \frac{\Phi^{-1}(\pi)}{\sqrt{1-\rho}} < 0 \text{ for } \pi < 1/2, \tag{38}$$

which yields the following useful relationship between a, δ and π :

$$a = \sqrt{1+\delta^2}\ \Phi^{-1}(\pi) . \tag{39}$$

Therefore, for a given value of $\pi < 1/2$, a and δ are negatively related and can not vary freely of one another.

Under the conditionally independent normal assumption,

$$\begin{pmatrix} \varepsilon_{i,t+1} \\ v_{ia} \\ v_{i\delta} \end{pmatrix} |f_t \backsim iidN\left(\mathbf{0}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega_{aa} & \omega_{a\delta} \\ 0 & \omega_{a\delta} & \omega_{\delta\delta} \end{pmatrix}\right),$$

then $\zeta_{i,t+1}\,|f_t \backsim iidN(0, 1+\omega_{aa}+\omega_{\delta\delta}f_t^2 - 2\omega_{a\delta}f_t)$, and hence as $N \to \infty$ the loss distribution can be simulated using[17]

$$x(f) = \Phi\left(\frac{a-\delta f}{\sqrt{1+\omega_{aa}+\omega_{\delta\delta}f^2 - 2\omega_{a\delta}f}}\right), \tag{40}$$

for random draws of $f \backsim N(0,1)$. Note that the asymptotic loss distribution is given by the distribution of x (the fraction of the portfolio lost) over $(0,1]$. Equation (40) is a key expression which we use below to analyze the impact of heterogeneity (or its neglect), manifested through non-zero values of $\omega_{aa}, \omega_{\delta\delta}$, and $\omega_{a\delta}$, on the loss distribution, especially its tail.

3.4.1 Heterogeneity of the Mean Returns and/or Default Thresholds

Consider first the case where the standardized factor loading is the same for all firms, namely $\delta_i = \delta$, $\forall i$, but allow for differences in a_i. This also imposes $\sigma_i^2 = \sigma^2$, $\forall i$, and implies the same pair-wise return correlation, ρ, across all firms. As a result, any variation in a_i is due to cross firm variation in $\lambda_i - \mu_i$, the difference between the default threshold and the mean return. It is unlikely that one would see differences in firm thresholds, perhaps due to management quality, but not in expected returns, so that variation in λ_i will likely be accompanied by variation in μ_i.

With that in mind, portfolio losses are

$$x = \Phi\left(\tilde{a} - \tilde{\delta} f\right),$$

where

$$\tilde{a} = \frac{a}{\sqrt{1+\omega_{aa}}},\ \tilde{\delta} = \frac{\delta}{\sqrt{1+\omega_{aa}}}. \tag{41}$$

[17] Here to simplify the exposition we have denoted the limit of $\ell_{N,t+1}$ by x, and have abstracted from the subscript t since f_t is serially uncorrelated.

In this case the CDF of x would have the same form as Vasicek's loss distribution, namely[18]

$$
\begin{aligned}
F_{\ell}(x) &= \Phi\left(\frac{\Phi^{-1}(x)-\tilde{a}}{\tilde{\delta}}\right), \qquad (42)\\
&= \Phi\left(\frac{\sqrt{(1+\omega_{aa})(1-\rho)}\Phi^{-1}(x)-\Phi^{-1}(\pi)}{\sqrt{\rho}}\right),
\end{aligned}
$$

where the second expression makes use of (38), (37) and (41). This clearly reduces to the CDF of the Vasicek's model for $\omega_{aa}=0$; see equation (A.4) in Appendix A.2.

It is also easily seen that in this case EL for the heterogeneous portfolio, denoted $\tilde{\pi}$, is given by

$$
\tilde{\pi}=E(x)=\Phi\left(\frac{\tilde{a}}{\sqrt{1+\tilde{\delta}^2}}\right)=\Phi\left(\frac{a}{\sqrt{1+\omega_{aa}+\delta^2}}\right)=\Phi\left(\frac{\Phi^{-1}(\pi)}{\sqrt{1+(1-\rho)\,\omega_{aa}}}\right), \qquad (43)
$$

which differs from the EL of the homogeneous portfolio. Since we are interested in losses for values of $\pi<1/2$, for which $\Phi^{-1}(\pi)<0$, we have

$$
\frac{\partial\tilde{\pi}}{\partial\omega_{aa}}=\phi\left(\frac{\Phi^{-1}(\pi)}{\sqrt{1+(1-\rho)\,\omega_{aa}}}\right)\ \frac{-\Phi^{-1}(\pi)(1-\rho)}{2\left(1+(1-\rho)\,\omega_{aa}\right)^{3/2}}\geq 0,\ \text{ for } \pi<1/2,
$$

and it readily follows that $\tilde{\pi}\geq\pi$, meaning EL is *under*estimated when $\omega_{aa}>0$ and this source of heterogeneity is neglected.

To derive the impact of ω_{aa} on unexpected loss, first without holding EL fixed, note that the pair-wise correlation of asset returns in this case is given by

$$
\tilde{\rho}=\frac{\delta^2}{1+\delta^2+\omega_{aa}}=\frac{\rho}{1+\omega_{aa}(1-\rho)},
$$

and using the results in Section 3.1 for N sufficiently large we have

$$
Var(x)=\tilde{\pi}\left(1-\tilde{\pi}\right)\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right),
$$

where

$$
\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right)=\frac{\Phi_2\left[\Phi^{-1}\left(\tilde{\pi}\right),\Phi^{-1}\left(\tilde{\pi}\right),\tilde{\rho}\right]-\tilde{\pi}^2}{\tilde{\pi}(1-\tilde{\pi})}, \qquad (44)
$$

and, as before in (15), $\Phi_2[\cdot]$ is the bivariate standard normal cumulative distribution function. Thus, if we allow EL to vary, the effect ω_{aa} on loss variance is

$$
\begin{aligned}
\frac{\partial Var(x)}{\partial\omega_{aa}} &= \frac{\partial Var(x)}{\partial\tilde{\pi}}\times\frac{\partial\tilde{\pi}}{\partial\omega_{aa}}+\frac{\partial Var(x)}{\partial\tilde{\rho}}\times\frac{\partial\tilde{\rho}}{\partial\omega_{aa}}\\
&= \left[\left(1-2\tilde{\pi}\right)\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right)+\tilde{\pi}\left(1-\tilde{\pi}\right)\left(\frac{\partial\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right)}{\partial\tilde{\pi}}\right)\right]\frac{\partial\tilde{\pi}}{\partial\omega_{aa}}+\tilde{\pi}\left(1-\tilde{\pi}\right)\left(\frac{\partial\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right)}{\partial\tilde{\rho}}\right)\frac{\partial\tilde{\rho}}{\partial\omega_{aa}}.
\end{aligned}
$$

The first term of this derivative is positive so long as $0\leq\tilde{\pi}<1/2$ and $\rho>0$ (and hence $\rho^{*}\left(\tilde{\pi},\tilde{\rho}\right)>0$).[19] However, the second term is negative since $\partial\tilde{\rho}/\partial\omega_{aa}<0$. Thus the net effect of heterogeneity in mean returns and/or default thresholds on portfolio loss variance is ambiguous.

[18] See Appendix A.2.

[19] Note that $\partial\tilde{\rho}^{*}/\partial\tilde{\pi}>0$.

3.4.2 Within and Between Type Heterogeneity

The paramter heterogeniety considered so far is relatively simple and can be viewed as within type heterogeneity, in the sense that differences across firms are random draws from the same common distribution. Under this set up, asymptotically as $N \to \infty$, the expected loss is invariant to the portfolio weights and it would not be possible to control the EL while experimenting with different degrees of heterogeneity as measured, for example, by different values of ω_{aa}. To control the EL we need to introduce an additional *systematic* source of heterogeneity. One possible approach would be to introduce firm types where for each type a_i's are draws from different distributions or from the same distribution but with different parameters. As an illustration, suppose the loan portfolio contains two types of firms, $\mathcal{H}$ and $\mathcal{L}$, with portfolio weights $w_{i\mathcal{H}}$, $i = 1, 2, ..., N_{\mathcal{H}}$, and $w_{i\mathcal{L}}$, for $i = 1, 2, ..., N_{\mathcal{L}}$, (such that $N = N_{\mathcal{H}} + N_{\mathcal{L}}$), and default probabilities, $\pi_{\mathcal{H}}$ and $\pi_{\mathcal{L}}$, respectively, with $0 < \pi_{\mathcal{L}} < \pi_{\mathcal{H}} < 1/2$. The differences in the default probabilities across the two types of firms could be due to differences in leverage or management quality, summarized, for instance, in a credit rating.

The portfolio loss in this case is given by

$$\ell_{N,t+1} = \sum_{i=1}^{N_{\mathcal{H}}} w_{i\mathcal{H}} I\left(a_{i\mathcal{H}} - \delta f_{t+1} - \varepsilon_{i\mathcal{H},t+1}\right) + \sum_{i=1}^{N_{\mathcal{L}}} w_{i\mathcal{L}} I\left(a_{i\mathcal{L}} - \delta f_{t+1} - \varepsilon_{i\mathcal{L},t+1}\right), \tag{45}$$

where $I(\cdot)$ is the indicator function as in (11),

$$a_{i\mathcal{H}} = a_{\mathcal{H}} + v_{i\mathcal{H}a}, \; a_{i\mathcal{L}} = a_{\mathcal{L}} + v_{i\mathcal{L}a}, \tag{46}$$

with $f \backsim N(0,1)$, $\varepsilon_{ik,t+1} \backsim N(0,1)$ and $v_{ika} \backsim N(0, \omega_{aa})$, for $k = \mathcal{H}, \mathcal{L}$. It is also assumed that $\varepsilon_{ik,t+1}$ and v_{ika} are independently distributed across all i and k.[20]

Let $w_{k,N_k} = \sum_{i=1}^{N_k} w_{ik}$, and $w_k = \lim_{N_k \to \infty} w_{k,N_k}$, $k = \mathcal{L}, \mathcal{H}$, where $w_{k,N_k} > 0$ for both finite N_k and as $N_k \to \infty$, so that $w_{\mathcal{H}}, w_{\mathcal{L}} > 0$, $w_{\mathcal{H},N_{\mathcal{H}}} + w_{\mathcal{L},N_{\mathcal{L}}} = 1$. Assuming that the granularity condition (16) holds for each firm type, then as $N_{\mathcal{H}}, N_{\mathcal{L}} \to \infty$ (the within-type portfolio must be large and granular to eliminate within type idiosyncratic risk), we have

$$x \,| f = w_{\mathcal{H}} \Phi\left(\tilde{a}_{\mathcal{H}} - \tilde{\delta} f\right) + w_{\mathcal{L}} \Phi\left(\tilde{a}_{\mathcal{L}} - \tilde{\delta} f\right), \tag{47}$$

where $\tilde{a}_k = a_k (1 + \omega_{aa})^{-1/2}$, for $k = \mathcal{H}, \mathcal{L}$, $w_{\mathcal{H}} + w_{\mathcal{L}} = 1$, and as before $\tilde{\delta} = \delta (1 + \omega_{aa})^{-1/2}$. Since $f \backsim N(0,1)$, it is now easily seen that

$$E(x) = \tilde{\pi} = w_{\mathcal{H}} \pi_{\mathcal{H}} + w_{\mathcal{L}} \pi_{\mathcal{L}},$$

where

$$\pi_k = \Phi\left(\frac{a_k}{\sqrt{1 + \omega_{aa} + \delta^2}}\right) = \Phi\left(\frac{a_k \sqrt{1 - \rho}}{\sqrt{1 + (1 - \rho)\, \omega_{aa}}}\right), \text{ for } k = \mathcal{H}, \mathcal{L},$$

[20] One could also allow for differences in the variances of v_{ika} across the types, $k = \mathcal{H}, \mathcal{L}$. But to keep the exposition simple here we are assuming that $Var(v_{ika}) = \omega_{aa}$.

and hence

$$a_k = \frac{\sqrt{1 + (1-\rho)\,\omega_{aa}}\ \Phi^{-1}(\pi_k)}{\sqrt{1-\rho}}, \text{ for } k = \mathcal{H}, \mathcal{L}.$$

To ensure the same expected losses under the homogeneous and heterogeneous cases we must have

$$\pi = w_{\mathcal{H}}\pi_{\mathcal{H}} + w_{\mathcal{L}}\pi_{\mathcal{L}}, \tag{48}$$

and this can be achieved, for given values of $\pi_{\mathcal{H}}$ and $\pi_{\mathcal{L}}$, by an appropriate choice of the portfolio weights on the types $\mathcal{L}$ and $\mathcal{H}$ (note that the granularity condition implies that changing the weights within type has no effect), so long as $\pi_{\mathcal{H}} \neq \pi_{\mathcal{L}}$, and $0 < \pi_k < 1$, for $k = \mathcal{H}, \mathcal{L}$.[21] Indeed both $w_{\mathcal{H}}$ and $w_{\mathcal{L}}$ must be positive, so long as $\pi_{\mathcal{H}} \neq \pi_{\mathcal{L}}$, to make the expected loss of the heterogeneous portfolio the same as for the homogeneous portfolio.

Using (45), and recalling the result in (15), we now have

$$V(x) = w_{\mathcal{H}}^2\left[F(\pi_{\mathcal{H}}, \pi_{\mathcal{H}}, \tilde{\rho}) - \pi_{\mathcal{H}}^2\right] + w_{\mathcal{L}}^2\left[F(\pi_{\mathcal{L}}, \pi_{\mathcal{L}}, \tilde{\rho}) - \pi_{\mathcal{L}}^2\right] + 2w_{\mathcal{H}}w_{\mathcal{L}}\left[F(\pi_{\mathcal{H}}, \pi_{\mathcal{L}}, \tilde{\rho}) - \pi_{\mathcal{H}}\pi_{\mathcal{L}}\right],$$

where

$$F(\pi_i, \pi_j, \tilde{\rho}) = \Phi_2\left[\Phi^{-1}(\pi_i), \Phi^{-1}(\pi_j), \tilde{\rho}\right].$$

Hence, under (48) the variance of the heterogeneous portfolio reduces to

$$V_{\text{het}}(x) = w_{\mathcal{H}}^2 F(\pi_{\mathcal{H}}, \pi_{\mathcal{H}}, \tilde{\rho}) + w_{\mathcal{L}}^2 F(\pi_{\mathcal{L}}, \pi_{\mathcal{L}}, \tilde{\rho}) + 2w_{\mathcal{H}}w_{\mathcal{L}}F(\pi_{\mathcal{H}}, \pi_{\mathcal{L}}, \tilde{\rho}) - \pi^2. \tag{49}$$

Furthermore, the variance of the associated homogeneous portfolio is given by

$$V_{\text{hom}}(x) = F(\pi, \pi, \rho) - \pi^2. \tag{50}$$

It is now easily established that so long as $w_{\mathcal{H}}$ (or $w_{\mathcal{L}}$) is set such that $\tilde{\pi} = \pi$, then for $\rho > 0$, $\omega_{aa} > 0$, and $a_{\mathcal{H}} \neq a_{\mathcal{L}}$, we have

$$V_{\text{hom}}(x) > V_{\text{het}}(x), \tag{51}$$

namely, the risk will be *over*estimated once the EL's of the two portfolios are equalized.

To prove this claim, note that since $\rho > \tilde{\rho}$, and $\partial F(\pi, \pi, \rho)/\partial\rho > 0$ (Vasicek 1998),

$$F(\pi, \pi, \rho) \geq F(\pi, \pi, \tilde{\rho}).$$

Therefore, to establish (51) it is sufficient to show that under $\pi = w_{\mathcal{H}}\pi_{\mathcal{H}} + w_{\mathcal{L}}\pi_{\mathcal{L}}$,

$$F(\pi, \pi, \tilde{\rho}) > w_{\mathcal{H}}^2 F(\pi_{\mathcal{H}}, \pi_{\mathcal{H}}, \tilde{\rho}) + w_{\mathcal{L}}^2 F(\pi_{\mathcal{L}}, \pi_{\mathcal{L}}, \tilde{\rho}) + 2w_{\mathcal{H}}w_{\mathcal{L}}F(\pi_{\mathcal{H}}, \pi_{\mathcal{L}}, \tilde{\rho}). \tag{52}$$

Consider now $F(x, y, \tilde{\rho})$ and note that $\partial^2 F(x, y, \tilde{\rho})/\partial x^2 < 0$,[22] and hence for given values of y and $\tilde{\rho}$, $F(x, y, \tilde{\rho})$ is concave in x and we have

$$F(\pi, \pi, \tilde{\rho}) = F(w_{\mathcal{H}}\pi_{\mathcal{H}} + w_{\mathcal{L}}\pi_{\mathcal{L}}, \pi, \tilde{\rho}) > w_{\mathcal{H}}F(\pi_{\mathcal{H}}, \pi, \tilde{\rho}) + w_{\mathcal{L}}F(\pi_{\mathcal{L}}, \pi, \tilde{\rho}). \tag{53}$$

[21] Note that the possibility of $\pi_{\mathcal{H}} = \pi_{\mathcal{L}}$ is ruled out only if $a_{\mathcal{H}} \neq a_{\mathcal{L}}$, which requires $a_{i\mathcal{H}}$ and $a_{i\mathcal{L}}$ to be draws from distributions with different means.

[22] A proof is provided in Appendix B.

Similarly, $\partial^2 F(x,y,\tilde{\rho})/\partial y^2 < 0$, and

$$\begin{aligned} F(\pi_{\mathcal{H}},\pi,\tilde{\rho}) &> w_{\mathcal{H}}F(\pi_{\mathcal{H}},\pi_{\mathcal{H}},\tilde{\rho}) + w_{\mathcal{L}}F(\pi_{\mathcal{H}},\pi_{\mathcal{L}},\tilde{\rho}), \\ F(\pi_{\mathcal{L}},\pi,\tilde{\rho}) &> w_{\mathcal{H}}F(\pi_{\mathcal{L}},\pi_{\mathcal{H}},\tilde{\rho}) + w_{\mathcal{L}}F(\pi_{\mathcal{L}},\pi_{\mathcal{L}},\tilde{\rho}). \end{aligned}$$

Using these results in (53), and noting that by symmetry $F(\pi_{\mathcal{H}},\pi_{\mathcal{L}},\tilde{\rho}) = F(\pi_{\mathcal{L}},\pi_{\mathcal{H}},\tilde{\rho})$, then (52) is readily established as required.

The above result is easily extended to portfolios containing more than two types of firms. Moreover, as the distance between $\pi_{\mathcal{L}}$ and $\pi_{\mathcal{H}}$ widens, the difference between the risks of the two portfolio types increases, suggesting that efficient credit portfolios should follow a "barbell" strategy combining exposures to very high quality credit with very low quality credits, so long as $\pi_k < 1/2$ for $k = \mathcal{H},\mathcal{L}$. As a result ignoring this type of heterogeneity would result in *over*estimation of risk when holding EL fixed.

3.4.3 Full Parameter Heterogeneity

The impact of allowing for full parameter heterogeneity in the multifactor case is discussed in Appendix A.3. For the single factor case, the analysis of allowing for non-zero values of ω_{aa}, $\omega_{\delta\delta}$, and $\omega_{a\delta}$ is easily carried out using (40) through random draws $f^{(r)} \backsim iidN(0,1)$ for $r = 1,2,...,R$. Given these simulated values one can readily compute UL, VaR, and other distributional characteristics as desired.

4 Illustrative Application: The Impact of Neglected Heterogeneity

In this section we consider different types of heterogeneity across firms and illustrate their effects on the resulting loss distribution by simulating losses for credit portfolios comprised of publicly traded U.S. firms. We also confirm that the predictions based on the random coefficient model, as set out in Section 3.4, match those obtained from more conventional simulation techniques. Finally, we are also interested in understanding which source of heterogeneity is the most important in affecting the shape of the loss distribution: the firm return process and associated factor loadings or the default threshold through information on distance to default or a credit rating.

4.1 Heterogeneity in Default Thresholds: Specification and Identification

We begin with a brief discussion of the specification and identification of the default thresholds. The probability of default for the i^{th} firm is given by (6), which we reproduce here for convenience:

$$\pi_{i,t+1} = \Phi\left(\frac{\lambda_{i,t+1} - \mu_i}{\sigma_{\xi_i}}\right).$$

This provides a functional relationship between a firm's equity returns (as characterized by μ_i and σ_{ξ_i}), its default threshold, $\lambda_{i,t+1}$, and the default probability, $\pi_{i,t+1}$. In the case of publicly traded companies, μ_i and σ_{ξ_i} can be consistently estimated from market returns based on historical data using either rolling or expanding observation windows. In general, however, $\lambda_{i,t+1}$ and $\pi_{i,t+1}$ can not be directly observed. One possibility would be to use balance sheet and other accounting data to estimate $\lambda_{i,t+1}$. This approach is taken up by Vassalou and Xing (2004) to cite an academic example, and KMV as an industry example, both of which use just the book value of debt (typically all short plus 1/2 of long term debt). But as argued in PSTW, such accounting information is likely to be noisy and might not be all that reliable due to information asymmetries and agency problems between managers, share-, and debtholders.[23] In addition to accounting data, other firm characteristics, such as firm age and perhaps size, as well as management quality could also be important in the determination of default thresholds, and most if not all of those characteristics typically go into a credit rating. In view of these measurement problems, PSTW propose an alternative estimation approach where firm-specific default thresholds are obtained using firm-specific credit ratings and historical default frequencies. These credit ratings could be either external, e.g. supplied by a rating agency, or internal from a bank's rating unit.

To be sure, neither our approach nor the results are predicated on the use of credit ratings per se, but rather on some summary measure of firm-specific default risk. PD point estimates, however derived, are very noisy, suggesting an averaging or grouping approach. This is effectively what a credit rating does, whether provided by an external rating agency or a bank-internal model. Moreover, since the bulk of a bank's lending portfolio is to privately held firms, typically only relatively coarse groupings are possible. See Hanson and Schuermann (2005) for a discussion on external ratings, and Trück and Ratchev (2005) on bank-internal ratings.

Broadly two identification schemes are possible, and they imply in turn assumptions about the distance to default, DD. One approach would be to impose the same default threshold for all firms of a given rating. Alternatively one could impose the same DD for all firms of a given rating, meaning

$$DD_{i,t+1} = \frac{\lambda_{i,t+1} - \bar{\mu}_i}{\bar{\sigma}_{\xi_i}} = DD_{\mathcal{R},t+1}, \tag{54}$$

for all firms i with rating $\mathcal{R}$, where $\bar{\mu}_i$ and $\bar{\sigma}_{\xi_i}$ are the *unconditional* estimates of μ_i and σ_{ξ_i} obtained using observations on firm-specific returns up to the end of period t. In this case the default threshold is different for every firm and can be computed using

$$\hat{\lambda}_{i,t+1} = \widehat{DD}_{\mathcal{R},t+1}\ \bar{\sigma}_{\xi_i} + \bar{\mu}_i, \text{ for } i \in \mathcal{R}_t, \tag{55}$$

[23] With this in mind, Duffie and Lando (2001) allow for the possibility of imperfect information about the firm's assets and default threshold in the context of a first-passage model. Their model is confirmed empirically in Yu (2005).

where

$$\widehat{DD}_{\mathcal{R},t+1} = \Phi^{-1}\left(\hat{\pi}_{\mathcal{R},t+1}\right), \tag{56}$$

$\Phi^{-1}(\cdot)$ is the inverse of the cumulative distribution function of the standard normal, and $\hat{\pi}_{\mathcal{R},t+1}$ is the observed default frequency of $\mathcal{R}$-rated firms.[24] This approach is analogous to the systematic heterogeneity by types discussed in the theory Section 3.4.2. Once again the estimated default thresholds, $\hat{\lambda}_{i,t+1}$, will be finite so long as $\hat{\pi}_{\mathcal{R},t+1} \neq \{0,1\}$.

The identification conditions can be summed up as follows: condition (54) imposes the same probability of default for each $\mathcal{R}-$rated firm, whereas the alternative strategy simply imposes that this needs to hold on average across $\mathcal{R}-$rated firms in the portfolio. Of the two, the assumption of the same distance-to-default seems more in line with the way credit ratings are established by the main rating agencies. First, the idea that firms with similar distances-to-default have similar probabilities of default is central to structural models of default. For instance, KMV makes use of a one-to-one mapping from DDs to EDFs (expected default frequencies). Second, rating agencies attempt to group firms according to their probability of default (subject possibly to some adjustments for differences in their expected loss given defaults), and in a structural model this is equivalent to grouping firms according to distance-to-default. In our empirical analysis we shall focus on the threshold estimates given by (55).[25]

The default threshold as specified in (55) incorporates equity market and credit rating information. Empirically, the right-hand-side of (55) is estimated on a rolling-window basis allowing for time variation in $\hat{\lambda}_{i,t+1}$.

4.2 Data and Portfolio Construction

We form credit portfolios of publicly traded U.S. firms at the end of each year from 1997 to 2002 and then simulate portfolio losses for the following year. Parameters are estimated recursively using 10-year (40-quarter) rolling windows. The simulations are out-of-sample in that the models, fitted over a ten-year sample, are used to simulate losses for the subsequent 11^{th} year. This recursive procedure allows us to explore the robustness of the results to possible time variation in the underlying parameters.

The loss simulations require an estimate of the probability of default for each firm. These are obtained at the level of the credit rating, $\mathcal{R}$, assigned to the firm by the two largest credit rating agencies: Moody's and S&P. In keeping with our overall empirical strategy, we estimate probabilities of default recursively for each grade using 10-year rolling windows of all firm rating histories from S&P. These probabilities are estimated using the time-homogeneous Markov or parametric duration

[24] Note that $\Phi^{-1}\left(\pi_{i,t+1}\right) < 0$ for $\pi_{i,t+1} < \frac{1}{2}$. In practice $\pi_{i,t+1}$ tends to be quite small.

[25] More detail as well as results using the same-threshold (λ) identifying assumption are given in Pesaran, Schuermann and Treutler (2005).

estimator discussed in Lando and Skødeberg (2002) and Jafry and Schuermann (2004). We impose a minimum annual probability of default (PD) of 0.001% or 0.1 basis points. Our estimated PDs for both $\mathcal{AAA}$ and $\mathcal{AA}$ fall below this minimum for all years.

In order to be selected for inclusion in one of our portfolios, a firm needs a) 10 years of consecutive quarterly equity returns in the CRSP database that match the rolling estimation window, and b) an active credit rating from either Moody's or S&P at the end of the window. In case both ratings are available the S&P rating is chosen.[26] For the first sample or cohort (which ends in 1997) we have 628 firms. At the end of the following year the portfolio is rebalanced, *retaining* surviving firms and *augmenting* the portfolio with new firms that have a rating at the end of that year, i.e. 1998, and also have 40 consecutive quarters of returns in the CRSP database.

The portfolio composition is adjusted annually, starting with 1998, to reflect defaults, upgrades and downgrades which may have occurred during the year. We also reallocate exposure annually to reflect changes in the distribution of ratings within the universe of rated U.S. firms. For example, at the end of 1997 $\mathcal{CCC}$-rated firms made up only 2.31% of all rated U.S. firms, but by year-end 2002 this proportion had risen to 5.89%. In Table 1 we show the average ratings distribution for 1997 and 2002. It becomes clear that there has been a systematic deterioration in average credit quality of rated U.S. firms over this period. In addition, estimated probabilities of default for non-investment grade ratings, and for $\mathcal{CCC}$ in particular, have risen noticeably over this period. As a result, the weighted average annual probability of default, $\hat{\pi}$, has increased from 1.60% for the year-end 1997 portfolio to 4.12% for the year-end 2002 portfolio. However, since firms choose whether or not to obtain a rating, this sample suffers from self-selection as does any sample which makes use of credit ratings. As a result it is unclear whether these patterns are reflective of the broader population of U.S. firms.

Using two-digit SIC codes we group firms into seven broad sectors to ensure a sufficient number of firms per sector. The sectors and percentage of firms by sector at year-end 1997 and 2002 are summarized in Table 2.

4.3 Model Specifications

Let $r_{ij,t+1}$ to be the return of firm i in sector j over the quarter t to $t+1$. Following the multi-factor return model given by (8), we employ the following return regressions adapted to our empirical applications:

$$r_{ij,t+1} = \alpha_{ij} + \boldsymbol{\beta}'_{ij}\mathbf{f}_{t+1} + u_{ij,t+1}, \tag{57}$$

where $\mathbf{f}_{t+1} \sim \left(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f\right)$, $\boldsymbol{\mu}_f$ is an $m \times 1$ vector of constants, and $\boldsymbol{\Sigma}_f$ is the covariance matrix of the common factors, also assumed fixed. In terms of the return parameters of (3) and (8), the expected

[26] Our decision rule is driven by the use of S&P ratings histories to compute the default probabilities $\pi_{\mathcal{R}}$.

return can be written as

$$\mu_{ij,t+1} = \alpha_{ij} + \boldsymbol{\beta}'_{ij}\boldsymbol{\mu}_f, \tag{58}$$

and the unexpected component as

$$\xi_{ij,t+1} = \boldsymbol{\beta}'_{ij}(\mathbf{f}_{t+1} - \boldsymbol{\mu}_f) + u_{ij,t+1}. \tag{59}$$

The total return variance is given by

$$\sigma^2_{\xi_{ij}} = \boldsymbol{\beta}'_{ij}\boldsymbol{\Sigma}_f\boldsymbol{\beta}_{ij} + \sigma^2_{ij}, \tag{60}$$

where σ^2_{ij} is the variance of the idiosyncratic component, $u_{ij,t+1}$. Note that (57) is not a forecasting equation. Moreover, since the factors are assumed to be serially uncorrelated here, there is no meaningful distinctions between conditional and unconditional returns and hence loss distributions. This is in contrast to the observable factor model presented in PSTW where the (global) factor structure is modeled as a vector autoregressive error correcting mechanism, thereby explicitly introducing serial correlation in the returns, making multi-period forecasts possible, conditional on the factor values at the end of the sample period.

Following a standard approach in the finance literature, we model firm returns using an unobserved components or factor approach, either single or multiple, with increasing degrees of heterogeneity. The empirical exercise involves a number of variations on the basic firm return equation given by (57) using market-cap weighted market returns $\bar{r}_{t+1}$ as proxies for two of the possible m common factors. Sector returns, $\bar{r}_{j,t+1}$, are computed in a similar fashion, namely using the market-cap weighted average of firm returns in that sector.[27]

The simplest model is the fully homogeneous return specification analogous to the one assumed by Vasicek:

$$r_{ij,t+1} = \alpha + \beta\bar{r}_{t+1} + u_{ij,t+1}, \tag{61}$$

with $u_{ij,t+1} \sim iidN(0, \sigma^2)$.

The second model tests the predictions made by theory in Section 3.4 by introducing heterogeneity in default thresholds by rating (Model II). The third model specification allows for full parameter heterogeneity where firm alphas, factor loadings and error variances are allowed to vary across firms. In the fourth specification we add an industry or sector factor so that each firm's return is regressed on $\bar{r}_{t+1}$ as well as on $\bar{r}_{j,t+1}$.

In the loss simulations we must impose conditional independence, but if we have failed to capture this dependence in the return model specifications, we will subsequently underestimate risk. With that in mind, the fifth and final model specification is the principal components (PC) model. We selected $\hat{m}$, the number of factors, using the IC_1 and IC_2 selection criteria proposed in Bai and Ng (2002), with the maximum number of factors set to 5. Both criteria yielded the same result of

[27] The weights for period $t+1$ are based on the average of the market capitalization at end of periods t and $t+1$.

two factors. The procedure was conducted for the 1997 cohort of firms, using the prior ten years of quarterly data. For tractability the number of factors was kept fixed for the subsequent cohort of firms, though the actual factors were, of course, re-estimated. Table 3 summarizes the five model specifications that we consider.

4.4 Return Regressions: Recursive Estimates

The return regression parameters, estimated recursively using a 10-year rolling window, are summarized in Table 4. We focus our discussion on the average pair-wise correlation of returns and the average pair-wise correlation of residuals as they map naturally into our loss modeling framework. The average pair-wise correlation of residuals is of particular interest since it gives an indication of how close a particular model is to conditional independence.

Starting with the results in Panel A of Table 4, we note that the in-sample average pair-wise correlation of quarterly returns for the first ten years (1988-1997) is 0.1933. The factor models generally do a good job of accounting for the cross-section correlation of returns, at least in-sample. The average pair-wise correlation of residuals for the whole portfolio is around 0.037 for the Vasicek and the single factor CAPM models. Adding an industry factor reduces that residual correlation to 0.022, and the PCA model by construction leaves almost no cross-section residual correlation. To be sure, there is no guarantee that this will hold out-of-sample. In-sample goodness of fit across models as measured by $\bar{R}^2$ (not reported in the table) range from 0.112 for the Vasicek to 0.208 for the sector CAPM to 0.220 for the PCA model.

Panels B through F in Table 4 show the recursive results using a 10-year rolling window for the next five ten-year periods. We note that average pair-wise cross-sectional correlations of firm returns remain at around 20% through 1999, but starting with the cohort of 1991-2000 (Panel D), the average correlation for the portfolio drops to 0.169. The sudden and substantial market reversals in the U.S. in March 2000 and the subsequent market declines probably play a strong role in explaining these results.[28] Similar patterns are also observed across the different models over the successive periods.

A further source of systematic heterogeneity across firms is the default probability or distance to default, captured for instance by the differences in their credit rating. It is reasonable to expect that the return processes of firms with a relatively high credit rating should on average exhibit a

[28] Throughout the analysis we have been assuming time invariant volatilities. While it is well known that high frequency (daily, weekly) firm returns exhibit volatility clustering, this effect tends to vanish as the data frequency declines due to temporal aggregation effects. Nonetheless, we conducted standard diagnostic tests for ARCH effects on all firm return regressions in the case of Model III and calculated the percentage of firm-specific return regressions in which the ARCH effects are significant at the 5% level. For most periods the percentage of firms with significant ARCH effects fell between 5 and 10%; the detailed results are available upon request from the authors. Overall the evidence is not sufficiently overwhelming to motivate ARCH modeling across all firms.

relatively low error variances and vice versa. This is indeed the case when we compare the average estimates of $\sigma_{ij}^2 = Var(u_{ij,t+1})$ across ratings. Table 5 shows averages of $\hat{\sigma}_{ij}$ for Models I (Vasicek) and II (Vasicek + Rating) where the estimates do not vary across firms, and for the CAPM where they do, based on the final 10-year cohort, 1993 – 2002. Similar results are obtained for other sample periods. The firm beta, $\bar{\beta}$, for Models I & II (recall this is a pooled estimate) is 0.867, and the firm error volatility, $\bar{\sigma}$, is 0.194. The next three rows show the average estimates by credit rating for Model III. Taking the last row first, $\bar{\sigma}_{\mathcal{R}}$ increases monotonically as we descend the credit spectrum, from 0.110 for $\mathcal{AAA}$ and $\mathcal{AA}$ firms, to 0.287 for $\mathcal{B}$-rated and 0.301 for $\mathcal{CCC}$-rated firms. No such clear pattern can be seen for firm betas, $\bar{\beta}_{\mathcal{R}}$. Thus credit ratings seem to sort firms by firm-specific risk but not by firm beta or factor loading, and in this way it might be reasonable to consider credit rating as being able to distinguish systematic differences in distance to default.

4.5 Impact of Heterogeneity on Credit Losses

We are now ready to generate loss distributions using the return and distance to default parameter estimates from Section 4.4. We begin by calibrating the analytical (asymptotic with $N \to \infty$) results presented in Section 3.4 assuming the parameter estimates are random draws from Gaussian processes. We then simulate the corresponding loss distributions using firm-specific estimates in the context of our finite-sized portfolio. In what follows we shall refer to the former as the "asymptotic approach," and the latter as the "finite sample approach." In both approaches we maintain the double-Gaussian assumption applied to the common factors, $\mathbf{f}_t$, and the idiosyncratic shocks, $\varepsilon_{ij,t+1}$. This allows us to compare these two approaches, and should help shed light on the validity of the Gaussian random coefficient model for the analysis of loss distributions. Moreover, as argued in Section 4.6, the asymptotic approach has important practical merit for portfolios where reliable firm-specific estimates of β and σ can not be obtained because of inadequate return histories, or because some of the firms in the portfolio might not be publicly traded companies.

4.5.1 Simulating Asymptotic Losses: Random Parameters Approach

For the homogeneous case the asymptotic results are given in Appendix A.1. In what follows we focus on unexpected losses given by the square root of (22), and various quantiles or VaRs.

The return process, using the notations introduced in Section 3.4, is defined by

$$r_{i,t+1} = \mu_i + \gamma_i f_{t+1} + \sigma_i \varepsilon_{i,t+1}, \quad \varepsilon_{i,t+1} \sim iidN(0,1), \quad f_{t+1} \sim N(0,1).$$

Given the idiosyncratic nature of the firm-specific shocks, $\varepsilon_{i,t+1}$, the common factor can be consistently estimated using the market return, denoted by $\bar{r}_{t+1}$ (see Pesaran 2006). This yields the familiar CAPM specification:

$$r_{i,t+1} = \alpha_i + \beta_i \bar{r}_{t+1} + \sigma_i \varepsilon_{i,t+1}, \quad \varepsilon_{i,t+1} \sim iidN(0,1), \quad \bar{r}_{t+1} \sim N(\bar{r}, \sigma_{\bar{r}}^2),$$

with the following relationships between the parameters of the two specifications

$$\mu_i = \alpha_i + \beta_i \bar{r}, \text{ and } \gamma_i = \beta_i \sigma_{\bar{r}}.$$

Thus firm i's default probability is

$$\pi_{i,t+1} = \Pr\left(r_{i,t+1} \leq \lambda_{i,t+1}\right) = \Phi\left(\frac{\lambda_{i,t+1} - \mu_i}{\sigma_{\xi_i}}\right),$$

where $\sigma_{\xi_i}^2 = \beta_i^2 \sigma_{\bar{r}}^2 + \sigma_i^2$.

Using the identification strategy of same distance-to-default by rating $\mathcal{R}$ implies the following rating-specific default thresholds:

$$\lambda_{i\mathcal{R}} = \sigma_{\xi_i} \Phi^{-1}(\pi_{\mathcal{R}}) + \mu_i. \tag{62}$$

A firm of type $\mathcal{R}$ defaults if $\mu_{i\mathcal{R}} + \gamma_{i\mathcal{R}} f_{t+1} + \sigma_{i\mathcal{R}} \varepsilon_{i,t+1} \leq \lambda_{i\mathcal{R}}$, or if $\delta_{i\mathcal{R}} f_{t+1} + \varepsilon_{i,t+1} \leq a_{i\mathcal{R}}$, where

$$\delta_{i\mathcal{R}} = \frac{\gamma_{i\mathcal{R}}}{\sigma_{i\mathcal{R}}} = \frac{\beta_{i\mathcal{R}} \sigma_{\bar{r}}}{\sigma_{i\mathcal{R}}}, \tag{63}$$

and

$$a_{i\mathcal{R}} = \sqrt{1 + \delta_{i\mathcal{R}}^2}\ \Phi^{-1}(\pi_{\mathcal{R}}). \tag{64}$$

The reduced form parameters $\delta_{i\mathcal{R}}$ and $a_{i\mathcal{R}}$ can now be estimated using the estimates of β_i and σ_i from the CAPM regressions (categorized by credit rating at the end of the sample), the default probability estimates by rating, $\hat{\pi}_{\mathcal{R}}$, given for 2002 in the last column of Table 1, and the unconditional mean and variance of the market return, $\bar{r}$ and $\hat{\sigma}_{\bar{r}}^2$. The parameters that enter the random parameters loss distribution can then be computed as sample moments by rating which we denote by $\hat{a}_{\mathcal{R}}$,$\hat{\delta}_{\mathcal{R}}$, $\hat{\omega}_{a_{\mathcal{R}}a_{\mathcal{R}}}$, $\hat{\omega}_{\delta_{\mathcal{R}}\delta_{\mathcal{R}}}$ and $\hat{\omega}_{a_{\mathcal{R}}\delta_{\mathcal{R}}}$, for $\mathcal{R} = 1, 2, ..., K$, where K is the number of credit rating categories (in our application 7), $\hat{\delta}_{i\mathcal{R}} = \hat{\beta}_{i\mathcal{R}} \hat{\sigma}_{\bar{r}} / \hat{\sigma}_{i\mathcal{R}}$, $\hat{a}_{i\mathcal{R}} = \sqrt{1 + \hat{\delta}_{i\mathcal{R}}^2}\ \Phi^{-1}(\hat{\pi}_{\mathcal{R}})$, $N_{\mathcal{R}}$ is the number of firms in the rating category $\mathcal{R}$ at the end of the sample (where the loss distribution is to be simulated). Note also that $\sum_{\mathcal{R}=1}^{K} N_{\mathcal{R}} = N$.

Using the above parameter estimates losses can be simulated as

$$x^{(r)} = \sum_{\mathcal{R}=1}^{K} w_{\mathcal{R}} \Phi\left(\frac{\hat{a}_{\mathcal{R}} - \hat{\delta}_{\mathcal{R}} f^{(r)}}{\sqrt{1 + \hat{\omega}_{a_{\mathcal{R}}a_{\mathcal{R}}} + \hat{\omega}_{\delta_{\mathcal{R}}\delta_{\mathcal{R}}} f^{(r)2} - 2\hat{\omega}_{a_{\mathcal{R}}\delta_{\mathcal{R}}} f^{(r)}}}\right), \tag{65}$$

where $w_{\mathcal{R}}$ is the weight of $\mathcal{R}$-rated firms in the portfolio ($\sum_{\mathcal{R}=1}^{K} w_{\mathcal{R}} = 1$), and $f^{(r)}$, $r = 1, 2, ..., R$ are random draws from $N(0, 1)$.[29] This formulation automatically sets the expected loss to be equal when introducing credit rating information since portfolio loss is just the weighted average of loss by rating.

[29] Clearly, draws from other distributions can also be considered.

Table 6 shows the relevant parameter estimates needed to generate losses in (65), by rating, for the last 10-year sample window, $1993-2002$. The estimate $\hat{a}_{\mathcal{R}}$ is just distance to default scaled by $\sqrt{1+\hat{\delta}_{i\mathcal{R}}^2}$. Since the average standardized factor loading, $\hat{\delta}_{\mathcal{R}}$, varies little across rating, the increase in $\hat{a}_{\mathcal{R}}$ as we descend the rating spectrum is driven by the increase in probability of default, $\hat{\pi}_{\mathcal{R}}$.

A similar effect is driving the declining within-rating parameter variance, $\hat{\omega}_{a_{\mathcal{R}}a_{\mathcal{R}}}$. When $\hat{\pi}_{\mathcal{R}}$ is high (e.g. for low ratings such as $\mathcal{B}$ and $\mathcal{CCC}$), $\Phi^{-1}(\hat{\pi}_{\mathcal{R}})$ is close to zero and the within-type variance of $\hat{a}_{i\mathcal{R}}, \hat{\omega}_{a_{\mathcal{R}}a_{\mathcal{R}}}$, shrinks. Note that this within-type variation is assumed to be purely random. We do not expect any systematic pattern with regard to factor loadings $\hat{\delta}_{\mathcal{R}}$ nor their within-rating dispersion, $\hat{\omega}_{\delta_{\mathcal{R}}\delta_{\mathcal{R}}}$. Finally, the correlation coefficient between the two sets of parameter estimates, denoted by $\hat{\rho}_{a_{\mathcal{R}}\delta_{\mathcal{R}}}$, is always negative, as expected from (64).

In Table 7 we present descriptive statistics of the loss distributions. We draw the reader's attention to the first set of columns labeled "Random Parameters/Asymptotic" which summarize the loss distributions using (65) with $R = 1,000,000$ for the three single-factor specifications, namely homogeneous Vasicek (Model I), Vasicek plus rating (Model II), and CAPM (Model III). To allow for easy comparison we hold EL fixed across models. For each model and each year, we show UL and two commonly reported quantiles (VaR), 99.0% and 99.9%.

We see clearly that allowing for parameter heterogeneity reduces risk, whether measured by UL or VaR. Taking for instance the first simulation year, 1998, we see that allowing for only heterogeneity in the a-parameter through rating-specific distance to default, UL drops by nearly 40%, from 1.54% to 0.94%, and 99.9% VaR drops by nearly a half from 12.19% to 6.88%, as expected from the theoretical results in Section 3.4.1. Allowing in addition for factor loading heterogeneity results in a further reduction of about one-third in UL to 0.65%, and of a tenth in 99.9% VaR to 6.21%. This basic pattern is repeated for subsequent years.

4.5.2 Simulating Finite Sample Losses

In this section we simulate the loss distribution for our finite sample portfolio using firm-specific parameter estimates. This exercise allows us to assess the performance of the simulated asymptotic loss distribution in predicting tail properties as compared to the finite sample results.

We simulate firm returns out-of-sample using (57), assuming that the systematic and idiosyncratic components are serially uncorrelated and independently distributed, meaning we impose conditional independence. The loss distributions for the different model specifications are then simulated using appropriate default thresholds, and assuming for simplicity no recovery in the event of default. All simulations are based on 500,000 replications, and the results are reported in the last set of columns labeled "Firm-Specific Parameters/Finite Sample" in Table 7. In addition to VaR we also calculated expected shortfall; the results are qualitatively no different, and so we report here only the VaR results. Broadly speaking, risk, measured either by UL or VaR, declines

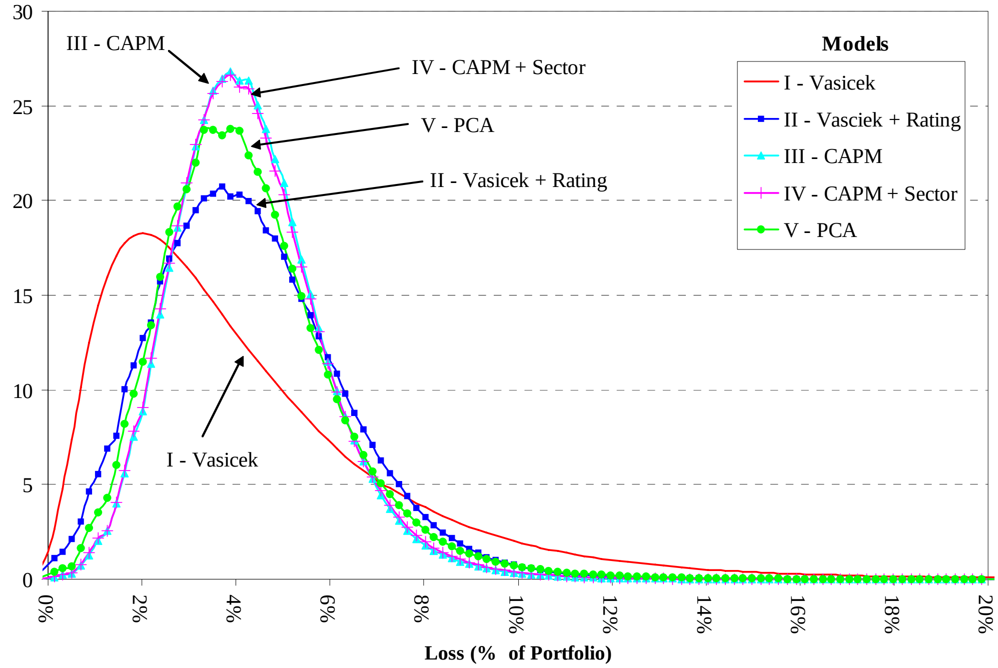

Figure 1: Simulated finite-sample loss densities using the firm-specific parameters approach (same EL, for 2003)

as model heterogeneity increases, and thus ignoring it would result in overestimation of risk.

Before embarking on a detailed model-by-model, year-by-year discussion of the loss simulations, it is helpful to consider Figure 1 to gain an overview of the results. In Figure 1 we show the loss densities for 2003 across the five different specifications. It is immediately apparent that the models are grouped into two sets. While the models differ in several ways, the main distinction between the two groups is the use of credit ratings. The more skewed density with the mode closer to the vertical axis is generated by the fully homogeneous Model I which does not make use of credit rating information while the others do. Indeed Model II adds only this source of information. Whatever other sources of heterogeneity may be important, an estimate of the probability of, or distance to, default, as provided by a credit rating, clearly has a significant influence on the overall shape of the loss density. Note that the rating serves as a group summary statistic of default risk and need not come from a rating agency. In a banking context it will likely be based on internal models.

We turn now to Table 7 where we report the finite sample loss simulations results for each of the six rolling windows. First, comparing the asymptotic and finite-sample results for Vasicek (Model I), we see that our finite portfolio is relatively close to an asymptotically diversified portfolio. For example, the portfolio in Panel A has 628 firms, or an effective number of 465 equal-sized exposures.[30] For this portfolio the simulated UL for the Vasicek model is 1.64%, only 10bps above the asymptotic result, and similarly for the two quantiles 99.0% (within 33bps) and 99.9% VaR (within 49bp).[31]

Comparing the first three models within Panel A, we note that the finite sample simulations confirm rather precisely the predictions made by the asymptotic theory, and Table 7 allows for easy side-by-side comparison. The fully homogeneous model of Vasicek (Model I) generates the most extreme losses and has the highest unexpected losses. Adding ratings information (Model II) results in a significant reduction in risk (while controlling for expected losses). If we start in Panel A, UL drops by 20% from 1.64% to 1.31% while 99.9% VaR is reduced by nearly 40% from 12.68% to 7.79%. Credit ratings seem to capture relevant firm-specific information, and this is useful even though the information is grouped together into just a few (seven) rating categories. Models III and IV allow for heterogeneous slopes (factor loadings) and firm-specific error variances, with Model IV also adding an industry return factor. UL falls another 30% from 1.31% in Model II to 1.14% in Model III, while 99.9% VaR declines another 10% from 7.79% to 7.09%. Thus, the ranking across these three models in the finite sample simulations are *exactly* as predicted using the random parameter approach in Section 4.5.1.

Adding an industry factor in Model IV results in a very small increase in risk from Model III. However, the distributions are extremely similar: UL is nearly the same, 1.14% vs. 1.16%. Finally, the principal components Model V generates UL results that are similar to Model II, which is Model I with ratings information, namely 1.40% vs. 1.31%. VaR, however, is higher. For instance, 99.9% VaR is 9.61%, compared to 7.09% for Model II. In this way Model V also generates tail losses which are higher than Models III and IV.

This upturn in risk may appear counter-intuitive – adding heterogeneity results not in risk reduction but in an increase. However, it is important to keep in mind that the out-of-sample loss simulations are performed under the maintained assumption of *conditional independence*. Recall from Table 4 that only Model V has an (in-sample) average pair-wise cross-sectional correlation of residuals which is effectively zero. All other models have some remaining correlation. Put differently, while Model V is conditionally independent on an in-sample basis, it seems that the others are not. So long as on an out-of-sample basis Model V is still closer to conditional independence than the others, and there is currently no way of verifying this, the other models will generate risk forecasts which are biased downward, meaning that risk would be underestimated since return correlations are underestimated. Measuring and evaluating out-of-sample conditional dependence

[30] If N is the number of obligors in our portfolio, each with exposure weight w_i which is randomly assigned, then $N^* = \left(\sum_{i=1}^{N} w_i^2\right)^{-1}$ is the equivalent number of equally weighted exposures.

[31] This comparison gives a clear indication of the granularity of our finite-sample portfolio since for Vasicek (Model I) $\hat{\omega}_{a_{\mathcal{R}}a_{\mathcal{R}}} = \hat{\omega}_{\delta_{\mathcal{R}}\delta_{\mathcal{R}}} = \hat{\omega}_{a_{\mathcal{R}}\delta_{\mathcal{R}}} = 0$ for all R. Thus, the differences between the two approaches for Model I reflect only the finite size of the portfolio in question.

is an important topic which requires further investigation.[32]

Returning to Figure 1, the loss densities are clearly very different. The Vasicek model has only three parameters (α, β, σ), and once credit rating information is included in Model II, the distributional shape changes dramatically. Indeed Model II yields a loss distribution which is remarkably similar to those generated by the fully heterogeneous model specification. Credit ratings seem indeed to be a useful and informative summary statistic for firm-level default risk.

Moving down the panels in Table 7 we notice that the portfolio is getting riskier over time; expected loss rises every year. If we compare value-at-risk, say at the 99.9%, for a model, say the one-factor CAPM model (Model III), we see that VaR increases from 7.09% in 1998 to 9.58% in 2000 to 11.68% in 2003.

4.6 Heterogeneity and the New Basel Capital Accord

Both sets of results indicate that neglecting firm heterogeneity can have significant impact on risk, and that a dominant source of heterogeneity is the difference in the probability of default across firms as summarized, for instance, in a credit rating. Credit ratings can be provided by an external rating agency or can be generated by the bank itself using internal rating models and tools. Indeed the New Basel Capital Accord allows banks, under Pillar 1 (prescribed minimum capital), to assign ratings to their obligors but does not allow them to compute their own factor loadings or return correlations. Those are fixed by the policy makers in the form of the risk weight function which assigns capital to credit exposures. It thus appears that the Basel Committee has, perhaps unwittingly, allowed banks to model the most important source of heterogeneity.

Nonetheless, full-blown economic capital models, which are limited only by available data, will play an important role in the New Basel Accord under Pillar 2 (supervisory discretion). For obligors that are publicly traded, the firm-specific parameters approach discussed in Section 4.5.2 would be such an example. However, the bulk of a bank's lending portfolio is to firms which are privately held, and so parameters such as factor loadings cannot be estimated at the firm level. It is still possible to estimate credit ratings using firm financials (these are reported to banks as part of the lending relationship). In this case the theoretical results based on the random parameters approach could be very helpful as the parameter means and their covariances could be estimated using data from publicly traded firms. Moreover, additional (systematic) heterogeneity with respect to country or industry sector could be accommodated in this way. Simply put, our theoretical results provide an easy way of incorporating fairly rich parameter heterogeneity for a portfolio of credit exposures when only limited data is available for parameter estimation.

[32] The absence of conditional indendence empirically, especially on an out-of-sample basis, we think has been neglected in the literature. Das, Duffie, Kapadia, and Saita (2005) also find significant remaining correlation, or default clustering, even after accounting for observable factors. They propose an unobserved factor approach they call "frailty" to absorb the remaining dependence, though this approach makes out-of-sample forecasting challenging.

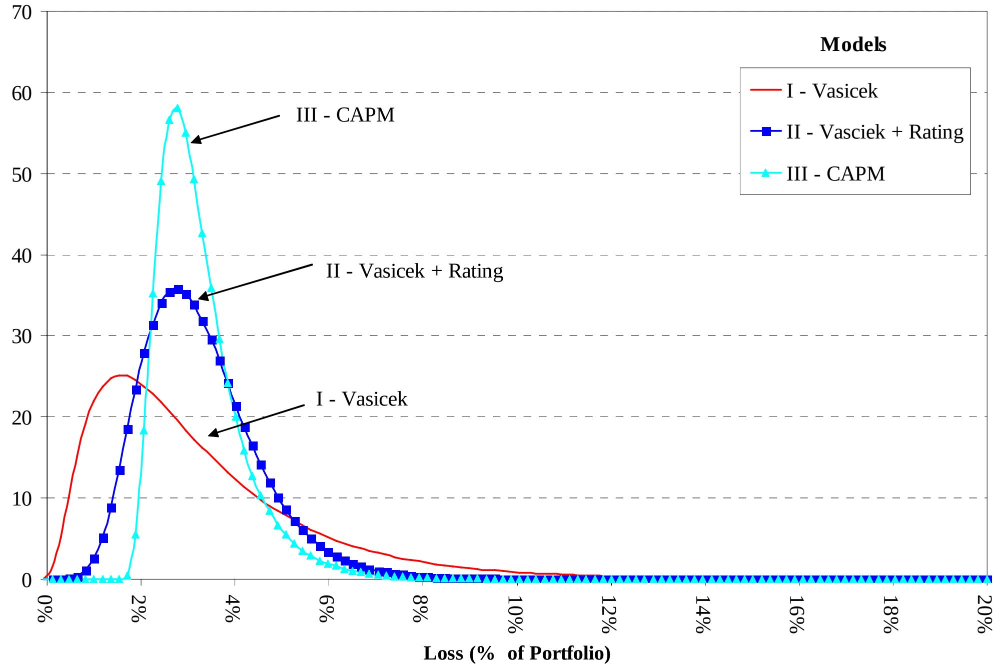

Figure 2: Calibrated (asymptotic) loss densities under random parameter framework (same EL, for 2003)

As an illustration, in Figure 2 we show the loss densities based on the random parameters approach as in Section 4.5.1 above using the estimates based on the last 10-year window (ending in 2002, thus simulating the loss distribution for 2003). The chart shows the densities for the three models (I: Vasicek; II: Vasicek plus rating; III: CAPM) while keeping EL fixed, and therefore the chart is comparable to the densities in Figure 1 which make use of the firm-specific parameter estimates rather than just their moments. Even with this limited information and the assumption of multivariate normality of the parameters made by the random coefficient framework, the densities in Figure 2 are remarkably close to those in Figure 1. The simple homogeneous Vasicek specification generates the most skewed and fat-tailed distribution. Allowing for ratings (Model II) has a significant effect on the shape of the loss density, as does adding factor loading heterogeneity (Model III). Returning now to Table 7, we note that the difference in VaR between the random parameters and firm-specific parameters approaches is quite small. For the simulation year used in Figures 1 and 2 (2003), the difference in 99.9% VaR is 3% for Model I, 7% for Model II and 8% for Model III. Similar differences are obtained for the other five years (not reported). However, the results for the homogenous Vasicek model show that the differences between the the finite sample and asymptotic outcomes are partly due to portfolio granularity. Thus, the results in Table 7 suggest that for large portfolios the differences between simulations based on the random parameters framework and those that make use of firm-specific parameter estimates are likely to be even smaller.

5 Concluding Remarks

In this paper we have considered a simple model of credit risk and derived the limit distribution of losses under different distributional assumptions regarding the structure of systematic and idiosyncratic risks and the nature of firm heterogeneity. The analytical and simulation results point to some interesting conclusions. Theory indicates that under the maintained assumption of conditional independence, meaning that all cross-firm dependence is captured by the systematic risk factors, if the firm parameters are heterogeneous but come from a common distribution, asymptotically (when the number of exposures, N, is sufficiently large) there is no scope for further risk reduction through active credit portfolio management. However, if firms are systematically different in that their parameters come from different distributions, as could be the case for firms from different sectors or countries, then further risk reduction is possible, even asymptotically, by changing the portfolio weights across types. In either case, *neglecting* parameter heterogeneity can lead to *under*estimation of expected losses. Then once expected losses are controlled for, neglecting parameter heterogeneity can lead to *over*estimation of risk, whether measured by unexpected loss or value-at-risk. Effectively the loss distribution is more skewed and fat-tailed when heterogeneity is ignored.

In light of these observations a natural question is: which sources of heterogeneity are most important from the perspective of portfolio losses? Here the answer seems clear: allowing for differences in the default threshold or probability of default (PD), measured for instance by a credit rating, is of first order importance in affecting the shape of the loss distribution. Including ratings heterogeneity alone results in a drop in loss volatility of 20%, and a drop of nearly 40% in 99.9% VaR, the VaR-level to which the New Basel Accord is calibrated. For policy makers and risk managers alike, this is good news. After all, an obligor PD, in the form of a rating, whether generated by a bank internally or provided by a rating agency externally, is one of the key parameters in the New Basel Accord which is allowed to vary. Indeed, early U.S. supervisory guidance indicates that banks must group their obligors into at least seven (non-default) grades, each with a unique PD (FRB 2003, p.201). Our results suggest that possibly finer differentiation or grouping along these lines may be fruitful, so long as it is properly done.[33]

When considering the return specification, flexible factor sensitivities also appear to be important, especially for capturing cross-firm dependence. If the maintained assumption of conditional independence is violated, i.e. if there remains cross-sectional dependence in the residuals from the return regressions, then risk will be *under*estimated. Thus proper specification of the return model is key by allowing for heterogenous factor loadings and the possible addition of industry return factors.

[33] See, for instance, Hanson and Schuermann (2005) for a discussion on PD estimation and their grouping into ratings.

A Appendix A: Limit Behavior of Credit Loss Distribution

A.1 Loss Densities under Homogeneous Parameters

In order to show how our approach relates to that of Vasicek, here we consider the homogeneous parameter case but do not require f_{t+1} and $\varepsilon_{i,t+1}$ to have Gaussian distributions. Since in the homogeneous case the multifactor model is equivalent to a single factor model, we consider scalar values for δ_i and f_{t+1} and denote them by δ and f_{t+1}, respectively. In this case we note that conditional on f_{t+1}, the random variables $z_{i,t+1}$ are identically and independently distributed as well as being integrable. (Recall that $|w_i z_{i,t+1}| \leq 1$ for all i and t.) Hence, conditional on f_{t+1} and as $N \to \infty$, we have

$$\ell_{N,t+1} \mid f_{t+1}, \mathcal{I}_t \overset{a.s.}{\to} F_\varepsilon \left(a - \delta f_{t+1}\right).$$

In the limit the probability density function of $\ell_{N,t+1} \mid I_t$ can be obtained from the probability density functions of f_{t+1} and $\varepsilon_{i,t+1}$, which we denote here by $f_f(\cdot)$ and $f_\varepsilon(\cdot)$, respectively. It will be helpful to write the loss density $f_\ell(\cdot)$ in terms of the systematic risk factor density $f_f(\cdot)$ and the standardized idiosyncratic shock density $f_\varepsilon(\cdot)$.

Therefore, conditional on I_t and denoting the limit of $\ell_{N,t+1}$ as $N \to \infty$, by ℓ_{t+1} we have (with probability 1)

$$\ell_{t+1} = F_\varepsilon \left(a - \delta f_{t+1}\right). \tag{A.1}$$

Now making use of standard results on transformation of probability densities, for $\delta \neq 0$ we have

$$f_\ell \left(\ell_{t+1} \mid \mathcal{I}_t\right) = \left| \frac{\partial F_\varepsilon \left(a - \delta f_{t+1}\right)}{\partial f_{t+1}} \right|^{-1} f_f \left(f_{t+1} \mid \mathcal{I}_t\right),$$

where f_{t+1} is given in terms of ℓ_{t+1}, via (A.1), namely

$$f_{t+1} = \frac{a - F_\varepsilon^{-1}\left(\ell_{t+1}\right)}{\delta},$$

and $\left|\partial F_\varepsilon \left(a - \delta f_{t+1}\right) / \partial f_{t+1}\right|$ is the Jacobian of the transformation which is given by

$$\frac{\partial F_\varepsilon \left(a - \delta f_{t+1}\right)}{\partial f_{t+1}} = -\delta f_\varepsilon \left(a - \delta f_{t+1}\right) = -\delta f_\varepsilon \left[F_\varepsilon^{-1}\left(\ell_{t+1}\right)\right].$$

Hence

$$f_\ell \left(\ell_{t+1} \mid \mathcal{I}_t\right) = \frac{f_f \left(\frac{a - F_\varepsilon^{-1}(\ell_{,t+1})}{\delta} \mid \mathcal{I}_t\right)}{|\delta| \, f_\varepsilon \left[F_\varepsilon^{-1}\left(\ell_{t+1}\right)\right]}, \text{ for } 0 < \ell_{t+1} \leq 1. \tag{A.2}$$

A.2 Relation to Vasicek's Loss Distribution

The above results provide a simple generalization of Vasicek's one-factor loss density distribution, derived in Vasicek (1991, 2002) and Gordy (2000), given by

$$f_\ell \left(x \mid \mathcal{I}_t\right) = \sqrt{\frac{1-\rho}{\rho}} \left\{ \frac{\phi \left[\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(\pi)}{\sqrt{\rho}}\right]}{\phi \left[\Phi^{-1}(x)\right]} \right\}, \text{ for } 0 < x \leq 1, \ \rho \neq 0, \tag{A.3}$$

where x denotes the fraction of the portfolio lost to defaults. The corresponding loss distribution is

$$F_\ell\left(x \mid \mathcal{I}_t\right) = \Phi\left[\frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(\pi)}{\sqrt{\rho}}\right]. \tag{A.4}$$

The density (A.2) reduces to (A.3) when $\mu_{ft} = 0$, and assuming that the innovations, f_{t+1} and $\varepsilon_{i,t+1}$ are both Gaussian. In this case

$$\begin{aligned} f_f\left(f_{t+1} \mid \mathcal{I}_t\right) &= \phi\left(f_{t+1}\right), \\ f_\varepsilon\left(\varepsilon_{i,t+1} \mid \mathcal{I}_t\right) &= \phi\left(\varepsilon_{i,t+1}\right), \ F_\varepsilon(\cdot) = \Phi(\cdot), \end{aligned}$$

and

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \frac{1}{|\delta|}\left\{\frac{\phi\left[\frac{a-\Phi^{-1}(x)}{\delta}\right]}{\phi\left[\Phi^{-1}(x)\right]}\right\}, \text{ for } 0 < x \leq 1, \ \ |\delta| \neq 0 \tag{A.5}$$

where we have used x for ℓ_{t+1}. Furthermore, in the homogeneous case

$$\delta = \sqrt{\frac{\rho}{1-\rho}}, \text{ for } \rho > 0, \tag{A.6}$$

and

$$\pi = \Phi\left(\frac{a}{\sqrt{1+\delta^2}}\right). \tag{A.7}$$

Hence

$$a = \frac{\Phi^{-1}(\pi)}{\sqrt{1-\rho}}. \tag{A.8}$$

Using (A.6) and (A.8) in (A.5) now yields Vasicek's loss density given by (A.3) (note that $\phi(x) = \phi(-x)$).

Under the double-Gaussian assumption, the distribution of $\delta f_{t+1} + \varepsilon_{t+1}$ (conditional on I_t) is also Gaussian with mean $\delta\mu_{ft}$ and variance $1 + \delta^2$. Therefore,

$$E\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \Phi\left(\frac{a - \delta\mu_{ft}}{\sqrt{1+\delta^2}}\right).$$

Using (A.8) and (A.6) the conditional mean loss can therefore be written as

$$E\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \Phi\left[\Phi^{-1}(\pi) - \sqrt{\rho}\mu_{ft}\right], \tag{A.9}$$

and reduces to π only when $\mu_{ft} = 0$. It is also interesting to note that under $\mu_{ft} \neq 0$, Vasicek's loss density and distributions become

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \sqrt{\frac{1-\rho}{\rho}}\left\{\frac{\phi\left[\sqrt{\frac{1-\rho}{\rho}}\Phi^{-1}(x) - \sqrt{\frac{1}{\rho}}\Phi^{-1}(\pi) + \mu_{ft}\right]}{\phi\left[\Phi^{-1}(x)\right]}\right\}, \text{ for } 0 < x \leq 1, \ \rho > 0. \tag{A.10}$$

For $\rho > 0$, the cumulative distribution function associated with this density is given by

$$F_\ell\left(x \mid \mathcal{I}_t\right) = \Phi\left(\sqrt{\frac{1-\rho}{\rho}}\Phi^{-1}(x) - \sqrt{\frac{1}{\rho}}\Phi^{-1}(\pi) + \mu_{ft}\right). \tag{A.11}$$

Also

$$\frac{\partial F_\ell\left(x \mid \mathcal{I}_t\right)}{\partial \mu_{ft}} = \phi\left(\sqrt{\frac{1-\rho}{\rho}}\Phi^{-1}(x) - \sqrt{\frac{1}{\rho}}\Phi^{-1}\left(\pi\right) + \mu_{ft}\right) > 0,$$

which shows that good news (a rise in μ_{ft}) reduces the probability of losses above a given thresholds, i.e. reduces value-at-risk, as to be expected.

A.3 Loss Densities under Full Parameter Heterogeneity

The impact on the loss distribution is considerably more complicated when we consider heterogeneity across the full set of parameters, including for instance the factor loadings. In this case, $\ell_{N,t+1}$, is given by (31): $\ell_{N,t+1} = \sum_{i=1}^{N} w_i I\left(a - \boldsymbol{\delta}'\mathbf{f}_{t+1} - \zeta_{i,t+1}\right)$. Since conditional on f_{t+1}, the composite errors, $\zeta_{i,t+1} = \varepsilon_{i,t+1} - v_{ia} + v'_{i\boldsymbol{\delta}} f_{t+1}$, are independently distributed across i, then

$$\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t \overset{a.s.}{\rightarrow} F_{\varkappa}\left(\frac{\boldsymbol{\theta}'\mathbf{g}_{t+1}}{\omega_{t+1}}\right),$$

where $a.s.$ denotes almost sure convergence, and as before $F_{\varkappa}\left(\cdot\right)$ denotes the cumulative distribution function of the standardized composite errors, $\varkappa_{i,t+1}$, defined by (35), $g_{t+1} = \left(1, -\mathbf{f}'_{t+1}\right)'$ and ω_{t+1} is given by (33). Once again the limiting distribution of credit loss depends on the conditional densities of $\zeta_{i,t+1}$ and f_{t+1}. For example, if $\left(\varepsilon_{i,t+1}, v_{ia}, \mathbf{v}'_{i\boldsymbol{\delta}}\right)$ follows a multivariate Gaussian distribution, then $\varkappa_{i,t+1} \mid f_{t+1}, I_t \sim iidN(0,1)$.

The probability density of the fraction of the portfolio lost, x, over the range (0,1), can be derived from the (conditional) joint probability density function assumed for the factors, f, by application of standard change-of-variable techniques to the non-linear transformation

$$x = F_{\varkappa}\left(\frac{a - \boldsymbol{\delta}'\mathbf{f}}{\sqrt{1 + \omega_{aa} - 2\boldsymbol{\omega}'_{a\delta}\mathbf{f} + \mathbf{f}'\Omega_{\delta\delta}\mathbf{f}}}\right). \tag{A.12}$$

For a general m factor set up analytical derivations are quite complicated and will not be attempted here. Instead, we consider the relatively simple case of a single factor model, where f is a scalar, f. Suppose $f = \psi(x)$ satisfies the transformation, (A.12), and note that

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \left|\psi'(x)\right| f_f\left[\psi(x) - \mu_{ft}\right], \text{ for } 0 < x \leq 1,$$

where $\left|\psi'(x)\right| = \left|x'(f)\right|^{-1}$. In other words, $\psi(x)$ is that value of the systematic factor f which generated loss of x. In the double-Gaussian case, for example, we have

$$\begin{aligned} x'(f) &= \left(\frac{f\left(\delta\omega_{a\delta} - a\omega_{\delta\delta}\right) + a\omega_{a\delta} - \delta\left(1 + \omega_{aa}\right)}{\left(1 + \omega_{aa} - 2\omega_{a\delta}f + \omega_{\delta\delta}f^2\right)^{3/2}}\right) \\ &\quad \times\phi\left(\frac{a - \delta f}{\sqrt{1 + \omega_{aa} - 2\omega_{a\delta}f + \omega_{\delta\delta}f^2}}\right). \end{aligned}$$

Hence

$$\left|\psi'(x)\right| = \left(\frac{1}{\phi\left[\Phi^{-1}(x)\right]}\right)\left|\frac{\left[1 + \omega_{aa} - 2\omega_{a\delta}\psi(x) + \omega_{\delta\delta}\psi^2(x)\right]^{3/2}}{\psi(x)\left(\delta\omega_{a\delta} - a\omega_{\delta\delta}\right) + a\omega_{a\delta} - \delta\left(1 + \omega_{aa}\right)}\right|.$$

and for $0 < x \leq 1$ we have

$$f_\ell(x \mid \mathcal{I}_t) = \left| \frac{\left[1+\omega_{aa} - 2\omega_{a\delta}\psi(x) + \omega_{\delta\delta}\psi^2(x)\right]^{3/2}}{\psi(x)\left(\delta\omega_{a\delta} - a\omega_{\delta\delta}\right) + a\omega_{a\delta} - \delta\left(1+\omega_{aa}\right)} \right| \left\{ \frac{\phi\left[\psi(x) - \mu_{ft}\right]}{\phi\left[\Phi^{-1}(x)\right]} \right\}, \tag{A.13}$$

This limiting loss distribution does not depend on the individual values of the portfolio weights, w_i, $i = 1, 2, ..., N$, so long as the granularity conditions in (16) are satisfied.

B Appendix B

Proposition 1 *Let $F(x_1, y_1, \rho) = \Phi_2(\Phi^{-1}(x_1), \Phi^{-1}(y_1), \rho)$. Then for $\rho > 0$, $\partial^2 F(x_1, y_1, \rho)/\partial^2 x_1 < 0$ and $\partial^2 F(x_1, y_1, \rho)/\partial^2 y_1 < 0$.*

Proof: By the symmetry of $F(x_1, y_1, \rho) = \Phi_2(\Phi^{-1}(x_1), \Phi^{-1}(y_1), \rho)$ in x_1 and y_1, it suffices to show that $\partial^2 F(x_1, y_1, \rho)/\partial x_1^2 < 0$. Let $G(x) = \Phi^{-1}(x)$ and note that

$$G'(x) = \Phi^{-1\prime}(x) = \frac{1}{\phi(\Phi^{-1}(x))} \text{ and } G''(x) = \Phi^{-1\prime\prime}(x) = \frac{\Phi^{-1}(x)}{\phi(\Phi^{-1}(x))^2}.$$

We have

$$\begin{aligned} F(x_1, y_1, \rho) &= \Phi_2(G(x_1), G(y_1), \rho), \\ &= \int_{-\infty}^{G(x_1)} \int_{-\infty}^{G(y_1)} \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho x p + y^2}{2(1-\rho^2)}\right) dy dx, \\ &= \int_{-\infty}^{G(x_1)} \int_{-\infty}^{G(y_1)} \frac{1}{\sqrt{1-\rho^2}} \phi\left(\frac{y - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x) dy dx. \end{aligned}$$

Hence

$$\begin{aligned} \frac{\partial F(x_1, y_1, \rho)}{\partial x_1} &= \int_{-\infty}^{G(y_1)} \frac{G'(x_1)}{\sqrt{1-\rho^2}} \phi\left(\frac{y - \rho G(x_1)}{\sqrt{1-\rho^2}}\right) \phi(G(x_1)) dy, \\ &= \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{G(y_1)} \phi\left(\frac{y - \rho G(x_1)}{\sqrt{1-\rho^2}}\right) dy > 0, \end{aligned}$$

where the second line follows from noting that

$$G'(x_1)\phi(G(x_1)) = \phi(\Phi^{-1}(x_1))/\phi(\Phi^{-1}(x_1)) = 1.$$

Thus, the second partial derivative is given by

$$\frac{\partial^2 F(x_1, y_1, \rho)}{\partial x_1^2} = \rho \frac{G'(x_1)}{1-\rho^2} \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{G(y_1)} \left[y - \rho G(x_1)\right] \phi\left(\frac{y - \rho G(x_1)}{\sqrt{1-\rho^2}}\right) dy$$

The integral in this term can be evaluated using standard results on the expectation of truncated normal variables, noting that $y\,|x_1 \sim N(\rho G(x_1), 1-\rho^2)$, we have

$$\begin{aligned}\frac{1}{\sqrt{1-\rho^2}}\int_{-\infty}^{G(y_1)} y\phi\left(\frac{y-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)dy &= E[y|y\leq G(y_1)]\Pr(y\leq G(y_1)) \\ &= \Phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)\left[\rho G(x_1)-\sqrt{1-\rho^2}\frac{\phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)}{\Phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)}\right].\end{aligned}$$

We also have

$$-\rho G(x_1)\frac{1}{\sqrt{1-\rho^2}}\int_{-\infty}^{G(y_1)}\phi\left(\frac{y-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)dy = -\rho G(x_1)\Phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)$$

Combining these terms we have

$$=\rho\frac{G'(x_1)}{1-\rho^2}\left[\Phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)\rho G(x_1)-\sqrt{1-\rho^2}\phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)-\rho G(x_1)\Phi\left(\frac{G(y_1)-\rho G(x_1)}{\sqrt{1-\rho^2}}\right)\right].$$

Thus, we conclude that for $\rho > 0$,

$$\frac{\partial^2 F(x_1,y_1,\rho)}{\partial x_1^2} = -\frac{\rho}{\sqrt{1-\rho^2}}\left(\frac{1}{\phi(\Phi^{-1}(x_1))}\right)\phi\left(\frac{\Phi^{-1}(y_1)-\rho\Phi^{-1}(x_1)}{\sqrt{1-\rho^2}}\right)\leq 0,$$

as desired.

References

[1] Altman, Edward I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance* 20, 589-609.

[2] Bai, Jushan and Serena Ng (2002). "Determining the Number of Factors in Approximate Factor Models." *Econometrica*, 70, 191-221.

[3] Basel Committee on Banking Supervision (2001). "The Internal Ratings Based Approach." Basel Committee on Banking Supervision, May; <http://www.bis.org/publ/bcbsca.htm>.

[4] Basel Committee on Banking Supervision (2005). "International Convergence of Capital Measurement and Capital Standards: A Revised Framework." <http://www.bis.org/publ/bcbs118.htm>, November.

[5] Black, Fischer, and John C. Cox (1976). "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions." *Journal of Finance* 31, 351-367.

[6] Broadie, Mark, Mikhail Chernov and Suresh Sundaresan (2005). "Optimal Corporate Securities Values in the Presence of Chapter 7 and Chapter 11." forthcoming, *Journal of Finance*.

[7] Chamberlain, Gary and Michael Rothschild (1983). "Arbitrage, Factor Structure and Mean-Variance Analysis on Large Asset Markets." *Econometrica* 51, 1281-1304.

[8] Collin-Dufresne, Pierre, Robert S. Goldstein, and Julien Hugonnier (2004). "A General Formula for Pricing Defaultable Claims." *Econometrica* 72, 1377–1407.

[9] Connor, Gregory and Robert A. Korajczyk (1995). "The Arbitrage Pricing Theory and Multifactor Models of Asset Returns." ch. 4 in Robert A. Jarrow, Vojislav Maksimovic, and W.T. Ziemba (eds.), *Handbooks in Operations Research and Management Science*, Vol. 9, Amsterdam, NL: Elsevier.

[10] Das, Sajiv. R., Darrell Duffie, Nikunj Kapadia, and Leandro Saita (2005), "Common Failings: How Corporate Defaults are Correlated." forthcoming, *Journal of Finance*.

[11] Davis, Mark and Violet Lo (2001). "Infectious Defaults." *Quantitative Finance* 1, 382-387.

[12] Driessen, Joost (2005). "Is default event risk priced in corporate bonds?" *Review of Financial Studies* 18, 165-195.

[13] Duffie, Darrell (2005). "Credit Risk Modeling with Affine Processes." *Journal of Banking & Finance* 29, 2751–2802.

[14] Duffie, Darrell and Nicolae Gârleanu (2001). "Risk and Valuation of Collateralized Debt Obligations." *Financial Analyst Journal*, January/February, 41-59.

[15] Duffie, Darrell and David Lando (2001). "Term Structures of Credit Spreads with Incomplete Accounting Information." *Econometrica* 69, 633-664.

[16] Duffie, Darrell, Leandro Saita, and Ke Wang (2005). "Multi-Period Corporate Default Prediction with Stochastic Covariates." forthcoming, *Journal of Financial Economics*.

[17] Duffie, Darrell and Kenneth J. Singleton (1999). "Simulating Correlated Defaults." Working Paper, Stanford Graduate School of Business.

[18] Duffie, Darrell and Kenneth J. Singleton (2003). *Credit Risk: Pricing, Measurement and Management.* Princeton, NJ: Princeton University Press.

[19] Embrechts, Paul, Alexander J. McNeil and D. Straumann (2001). "Correlation and Dependency in Risk Management: Properties and Pitfalls." pp.176-223 in M.A.H. Dempster and H.K Moffat (eds.) *Risk Management: Value at Risk and Beyond*, Cambridge, UK: Cambridge University Press.

[20] Federal Reserve Board, 2003, "Supervisory Guidance on Internal Ratings-Based Systems for Corporate Credit," Attachment 2 in http://www.federalreserve.gov/boarddocs/meetings/2003/20030711/attachment.pdf.

[21] Fernholz, Robert, Ioannis Karatzas and Constantinos Kardaras (2003). "Diversity and Relative Arbitrage in Financial Markets." Department of Mathematics and Statistics, Columbia University.

[22] Frey, Rüdiger and Alexander J. McNeil (2003). "Dependent Defaults in Models of Portfolio Credit Risk." *Journal of Risk* 6, 59-92.

[23] Garbade, Kenneth (2001). *Pricing Corporate Securities as Contingent Claims*, Cambridge, MA: MIT Press.

[24] Giesecke, Kai and Stefan Weber (2004). "Cyclical Correlations, Credit Contagion and Portfolio Losses." *Journal of Banking & Finance* 28, 3009-3036.

[25] Gordy, Michael B. (2000). "A Comparative Anatomy of Credit Risk Models." *Journal of Banking & Finance* 24, 119-149.

[26] Gordy, Michael B. (2003). "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules." *Journal of Financial Intermediation* 12, 199-232.

[27] Gordy, Michael B. (2004). "Granularity Adjustment in Portfolio Credit Risk Measurement." in Giorgio P. Szegö (ed.), *Risk Measures for the 21st Century.* New York, NY: John Wiley & Sons.

[28] Gupton, Gregory M., Christopher C. Finger and Mickey Bhatia (1997). CreditMetrics^{TM} – Technical Document, this version: April 2, 1997. J.P. Morgan, New York.

[29] Hanson, Samuel G. and Til Schuermann (2005). "Confidence Intervals for Probabilities of Default." forthcoming, *Journal of Banking & Finance.*

[30] Hillegeist, Stephen A., Elizabeth K. Keating, Donald P. Cram and Kyle G. Lundsted (2004). "Assessing the Probability of Bankruptcy." *Review of Accounting Studies* 9, 5-34.

[31] Hull, John C. and Alan White (2006), "Valuing Credit Derivatives Using an Implied Copula Approach," mimeo, Rotman School of Management, University of Toronto.

[32] Jafry, Yusuf and Til Schuermann (2004). "Measurement, Estimation and Comparison of Credit Migration Matrices." *Journal of Banking & Finance* 28, 2603-2639.

[33] Jarrow, Robert A., David Lando, and Fan Yu (2005). "Default Risk and Diversification: Theory and Empirical Implications." *Mathematical Finance* 15, 1-26.

[34] Koyluoglu, H. Ugur, and Andrew Hickman (1998). "Reconcilable Differences." *Risk* 11, October, 56-64

[35] Lando, David (1998). "On Cox Processes and Credit Risky Securities." *Review of Derivatives Research* 2, 88-120.

[36] Lando, David (2004). *Credit Risk Modeling: Theory and Applications.* Princeton, NJ: Princeton University Press.

[37] Lando, David and Torben Skødeberg (2002). "Analyzing Ratings Transitions and Rating Drift with Continuous Observations." *Journal of Banking & Finance* 26, 423-444.

[38] Leland, Hayne E. and Klaus Bjerre Toft (1996). "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads." Journal of Finance 51, 987-1019.

[39] Lennox, Clive (1999). "Identifying Failing Companies: A Re-Evaluation of the Logit, Probit and DA Approaches." *Journal of Economics and Business* 51, 347-364.

[40] Li, David X. (2000). "Default Correlation: A Copula Approach." *Journal of Fixed Income*, March, 43-54.

[41] Lucas, André, Pieter Klaasen, Peter Spreij, and Stefan Straetmans (2001). "An Analytic Approach to Credit Risk of Large Corporate Bond and Loan Portfolios." *Journal of Banking & Finance* 25, 1635-1664.

[42] Merton, Robert C. (1974). "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29, 449-470.

[43] Pesaran, M.Hashem (2006). "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure.", *Econometrica,* July issue.

[44] Pesaran, M.Hashem, Til Schuermann, and Björn-Jakob Treutler (2005). "Global Business Cycles and Credit Risk." forthcoming in Mark Carey and René Stulz (eds.), *Risks of Financial Institutions*, Chicago, IL: University of Chicago Press. Available as NBER Working Paper w11493.

[45] Pesaran, M. Hashem, Til Schuermann, Björn-Jakob Treutler, and Scott M. Weiner (2005). "Macroeconomic Dynamics and Credit Risk: A Global Perspective." forthcoming, *Journal of Money, Credit and Banking.*

[46] Ross, Stephen A. (1976). "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13, 341-360.

[47] Saunders, Anthony and Linda Allen (2002). *Credit Risk Measurement- New Approaches to Value at Risk and Other Paradigms.* 2^{nd} Ed., New York: John Wiley & Sons.

[48] Schönbucher, Philipp J. (1998). "The Term Structure of Defaultable Bond Prices." *Review of Derivatives Research* 2, 161-192.

[49] Schönbucher, Philipp J. (2002). "Taken to the Limit: Simple and Not-so-Simple Loan Loss Distributions." Working Paper, Department of Statistics, Bonn Univ.

[50] Schönbucher, Philipp J. (2003), *Credit Derivatives Pricing Models: Models, Pricing and Implementation*, New York, NY: John Wiley & Sons.

[51] Shumway, Tyler (2001). "Forecasting Bankruptcy more Accurately: A Simple Hazard Model." *Journal of Business* 74, 101-124.

[52] Silverman, Bernard W. (1986). *Density Estimation for Statistics and Data Analysis.* New York, NY: Chapman & Hall.

[53] Trück, Stefan and Svetlozar T. Rachev (2005). "Credit Portfolio Risk and PD Confidence Sets through the Business Cycle," University of Karlsruhe Working Paper.

[54] Vasicek, Oldrich (1987). "Probability of Loss on Loan Portfolio." KMV Corp.; available at www.kmv.com.

[55] Vasicek, Oldrich (1991). "Limiting Loan Loss Distribution." KMV Corp.; available at www.kmv.com.

[56] Vasicek, Oldrich (1998). "A Series Expansion for the Bivariate Normal Integral." *Journal of Computational Finance* 1, 5-10.

[57] Vasicek, Oldrich (2002). "Loan Portfolio Value." Risk 15, 160-162.

[58] Vassalou, Maria, and Yuhang Xing (2004). "Default Risk in Equity Returns." *Journal of Finance* 59, 831-868.

[59] Wilson, Thomas (1997a). "Portfolio Credit Risk, Part I." *Risk* 10, September, 111-117.

[60] Wilson, Thomas (1997b). "Portfolio Credit Risk, Part II." *Risk* 10, October, 56-61.

[61] Zhou, Chunsheng (2001). "An Analysis of Default Correlations and Multiple Defaults." *Review of Financial Studies* 14, 555-576.

Table 1
Ratings Distributions and Probabilities of Default

	1997		2002	
Credit Rating	$w_R(\%)$	$\hat{\pi}_R(\%)$	$w_R(\%)$	$\hat{\pi}_R(\%)$
$\mathcal{AAA}$	2.86	0.001	2.37	0.001
$\mathcal{AA}$	10.81	0.001	9.25	0.001
$\mathcal{A}$	25.61	0.005	21.49	0.006
$\mathcal{BBB}$	22.33	0.064	25.67	0.106
$\mathcal{BB}$	16.3	0.481	15.72	0.630
$\mathcal{B}$	19.79	3.343	19.62	5.429
$\mathcal{CCC}$	2.31	36.487	5.89	49.776
Portfolio		**1.60**		**4.12**

Note: The table presents the distribution of firms by rating for year-end 1997 and 2002. These distributions are calculated by taking the average of the distribution for Moody's and the distribution for Standard and Poor's. We construct our portfolios so that the exposure weights are consistent with the observed rating distribution. The final column, $\hat{\pi}_R(\%)$, contains the estimated annual probabilities of default (PD) that are used in the simulation exercises. These PDs are estimated using the time-homogeneous Markov or parametric duration estimator discussed in Lando and Skødeberg (2002) and Jafry and Schuermann (2004). A minimum annual PD of 0.001% or 0.1 basis points is imposed.

Table 2
Industry Breakdown

Industry	**1997**	**2002**
Agriculture, Mining & Construction	5.3	5.7
Communication, Electric & Gas	16.7	12.8
Durable Manufacturing	22.1	23.3
Finance, Insurance & Real Estate	23.1	21.7
Non-durable Manufacturing	18.2	18.7
Service	4.8	7.2
Wholesale & Retail Trade	9.9	10.7
Total	**100.0**	**100.0**

Note: The table presents the distribution of firms by industry group as of year-end 1997 and 2002

Table 3
Specifications of Return Equations

Model		Return Specification
I	Vasicek	$r_{ij,t+1} = \alpha_c + \beta \overline{r}_{t+1} + u_{ij,t+1}$
II	Vasicek + Rating	$r_{ij,t+1} = \alpha + \beta \overline{r}_{t+1} + u_{ij,t+1}$
III	CAPM	$r_{ij,t+1} = \alpha_{ij} + \beta_{ij} \overline{r}_{ct+1} + u_{ij,t+1}$
IV	CAPM + Sector	$r_{ij,t+1} = \alpha_{ij} + \beta_{1,ij} \overline{r}_{t+1} + \beta_{2,ij} \overline{r}_{j,t+1} + u_{ij,t+1}$
V	PCA	$r_{ij,t+1} = \alpha_{ij} + \boldsymbol{\beta}_{ij}' \mathbf{f}_{t+1} + u_{ij,t+1}$

Note: $r_{ij,t+1}$ denotes the return of firm *i* in sector *j* over the quarter *t* to *t*+1. In Models I through IV, $\overline{r}_{t+1}$ denotes the market-cap weighted over the quarter *t* to *t*+1 and $\overline{r}_{j,t+1}$ is the market-cap weighted return for sector *j* over the same period. For Model I there is a single default threshold for all firms. For Models II to V there is one default threshold per rating.

Table 4
Average Pair-wise Correlation of Returns and In-sample Residuals Based on Ten-Year Rolling Windows

		Avg. Pair-wise Correlation of Returns	Model Specifications		Avg. Pair-wise Correlation of Residuals
Panel A	**1988-1997**	0.1933	I	Vasicek	0.0365
	# of firms	628	III	CAPM	0.0374
			IV	CAPM + Sector	0.0221
			V	PCA	0.0005
Panel B	**1989-1998**	0.2114	I	Vasicek	0.0440
	# of firms	633	III	CAPM	0.0456
			IV	CAPM + Sector	0.0324
			V	PCA	0.0004
Panel C	**1990-1999**	0.2237	I	Vasicek	0.0731
	# of firms	613	III	CAPM	0.0778
			IV	CAPM + Sector	0.0621
			V	PCA	-0.0001
Panel D	**1991-2000**	0.1691	I	Vasicek	0.0783
	# of firms	588	III	CAPM	0.0821
			IV	CAPM + Sector	0.0615
			V	PCA	0.0008
Panel E	**1992-2001**	0.1633	I	Vasicek	0.0740
	# of firms	585	III	CAPM	0.0772
			IV	CAPM + Sector	0.0624
			V	PCA	-0.0003
Panel F	**1993-2002**	0.1999	I	Vasicek	0.0772
	# of firms	600	III	CAPM	0.0811
			IV	CAPM + Sector	0.0658
			V	PCA	-0.0009

Note: This table presents the results of recursive estimation of return equations using quarterly return data. All estimation results are calculated using a 40-quarter rolling window. Portfolio determination and sample construction are discussed in Section 4.2. Specification of the return models is discussed in Section 4.3 (see Table 3 for further detail). The data source for returns is CRSP.

Table 5
Parameter Estimates by Credit Ratings based on Return Regressions
1993-2002

Model	Parameter	$\mathcal{AAA}$ / $\mathcal{AA}$	$\mathcal{A}$	$\mathcal{BBB}$	$\mathcal{BB}$	$\mathcal{B}$	$\mathcal{CCC}$
Vasicek /	$\bar{\beta}$	0.867	0.867	0.867	0.867	0.867	0.867
Vasicek + Rating	$\bar{\sigma}$	0.194	0.194	0.194	0.194	0.194	0.194
CAPM	$\bar{\beta}_{\mathcal{R}}$	0.847	0.821	0.757	0.984	1.161	0.362
	$\bar{\sigma}_{\mathcal{R}}$	0.110	0.128	0.149	0.219	0.287	0.301

Note: Averages of estimated parameters from Models I and III; see Table 3. Note that the return specification for Model II is the same as for Model I. $\bar{\beta}$ and $\bar{\sigma}$ are the pooled estimates of return factor loading ("beta") and firm error variance, respectively, and are thus invariant across ratings. In the CAPM model these parameters are estimated separately for each firm, so that $\bar{\beta}_{\mathcal{R}}$ and $\bar{\sigma}_{\mathcal{R}}$ are averages by rating.

Table 6
Parameter Means Used in Asymptotic Loss Distribution
1993-2002

Parameter	$\mathcal{AAA}$ / $\mathcal{AA}$	$\mathcal{A}$	$\mathcal{BBB}$	$\mathcal{BB}$	$\mathcal{B}$	$\mathcal{CCC}$
$\hat{\alpha}_{\mathcal{R}}$	-5.451	-4.524	-3.392	-2.738	-1.726	-0.006
$\hat{\delta}_{\mathcal{R}}$	0.607	0.489	0.393	0.365	0.317	0.132
$\omega_{a_{\mathcal{R}}a_{\mathcal{R}}}$	0.528	0.253	0.078	0.059	0.021	<0.001
$\omega_{\delta_{\mathcal{R}}\delta_{\mathcal{R}}}$	0.089	0.076	0.052	0.062	0.056	0.064
$\rho_{a_{\mathcal{R}}\delta_{\mathcal{R}}}$	-0.983	-0.968	-0.949	-0.952	-0.941	-0.568

Note: Averages (by rating) for reduced form parameters for CAPM model from Table 5. $\mathcal{AAA}$ and $\mathcal{AA}$ are grouped since their default probabilities are both assigned the minimum of 0.01bp. For details on how the parameters are computed, please see Section 4.5.1.

Table 7
Out-of-Sample Simulated Annual Losses Based on 10-Year Rolling Return Regressions

	Simulation Year	Using Sample	Model	Random Parameters, Asymptotic: UL	Random Parameters, Asymptotic: Value-at-Risk 99.0%	Random Parameters, Asymptotic: Value-at-Risk 99.9%	Firm-specific Parameters, Finite Sample: UL	Firm-specific Parameters, Finite Sample: Value-at-Risk 99.0%	Firm-specific Parameters, Finite Sample: Value-at-Risk 99.9%
Panel A	**1998**	**1988-1997**	I Vasicek	1.54%	7.46%	12.19%	1.64%	7.79%	12.68%
		EL= 1.60%	II Vasicek + Rating	0.94%	4.79%	6.88%	1.31%	5.55%	7.79%
			III CAPM	0.65%	4.03%	6.21%	1.14%	4.92%	7.09%
			IV CAPM + Sector	-	-	-	1.16%	5.01%	7.36%
			V PCA	-	-	-	1.40%	6.14%	9.61%
Panel B	**1999**	**1989-1998**	I Vasicek	2.03%	9.83%	16.00%	2.13%	10.07%	16.35%
		EL= 2.04%	II Vasicek + Rating	1.21%	6.13%	8.80%	1.52%	6.84%	9.61%
			III CAPM	0.77%	4.92%	7.42%	1.23%	5.73%	8.27%
			IV CAPM + Sector	-	-	-	1.29%	6.00%	8.54%
			V PCA	-	-	-	1.53%	7.19%	10.99%
Panel C	**2000**	**1990-1999**	I Vasicek	2.48%	11.96%	18.80%	2.60%	12.33%	19.21%
		EL= 2.70%	II Vasicek + Rating	1.42%	7.32%	10.13%	1.77%	8.09%	11.04%
			III CAPM	0.91%	6.01%	8.61%	1.42%	6.90%	9.58%
			IV CAPM + Sector	-	-	-	1.51%	7.17%	9.87%
			V PCA	-	-	-	1.78%	8.54%	12.84%
Panel D	**2001**	**1991-2000**	I Vasicek	2.05%	10.04%	14.68%	2.20%	10.48%	15.21%
		EL= 2.94%	II Vasicek + Rating	1.16%	6.46%	8.34%	1.59%	7.35%	9.45%
			III CAPM	1.11%	6.55%	8.75%	1.55%	7.38%	9.75%
			IV CAPM + Sector	-	-	-	1.64%	7.64%	10.18%
			V PCA	-	-	-	1.88%	8.72%	12.29%
Panel E	**2002**	**1992-2001**	I Vasicek	2.32%	11.40%	16.38%	2.47%	11.85%	16.93%
		EL= 3.48%	II Vasicek + Rating	1.26%	7.21%	9.11%	1.73%	8.09%	10.11%
			III CAPM	0.99%	6.72%	8.79%	1.56%	7.66%	9.77%
			IV CAPM + Sector	-	-	-	1.61%	7.88%	10.20%
			V PCA	-	-	-	1.86%	8.98%	12.28%
Panel F	**2003**	**1993-2002**	I Vasicek	3.11%	15.04%	22.07%	3.24%	15.43%	22.81%
		EL= 4.12%	II Vasicek + Rating	1.63%	8.99%	11.54%	1.99%	9.67%	12.34%
			III CAPM	1.06%	7.92%	10.72%	1.60%	8.77%	11.68%
			IV CAPM + Sector	-	-	-	1.64%	8.94%	11.85%
			V PCA	-	-	-	1.90%	10.07%	13.95%

Note: This table presents results for simulated out-of-sample annual loss distributions. Model specifications, including the return regressions and determination of default thresholds, are discussed in Section 4.3 (see Table 3 for more detail on the model specifications). Simulations are carried out using 1,000,000 replications for the analytical simulations and 500,000 replications for the finite sample replications. For each year all models are calibrated to have the same expect loss given by $\hat{\pi}$. For each simulation, the table reports the standard deviation of losses (denoted Unexpected Losses - UL) as well as the 99.0% and 99.9% quantiles of the distribution (denoted Value-at-Risk).