# Scope for Credit Risk Diversification

Samuel Hanson, M. Hashem Pesaran
and Til Schuermann

May 2005

CWPE 0519

**Not to be quoted without permission**

# Scope for Credit Risk Diversification[*]

Samuel Hanson[†]
Federal Reserve Bank of New York

M. Hashem Pesaran
University of Cambridge and USC

Til Schuermann[†]
Federal Reserve Bank of New York and Wharton Financial Institutions Center

February 2005

## Abstract

This paper considers a simple model of credit risk and derives the limit distribution of losses under different assumptions regarding the structure of systematic risk and the nature of exposure or firm heterogeneity. We derive fat-tailed correlated loss distributions arising from Gaussian (i.e. non-fat-tailed) risk factors and explore the potential for (and limit of) risk diversification. Where possible the results are generalized to non-Gaussian distributions. The theoretical results indicate that if the firm parameters are heterogeneous but come from a common distribution, for sufficiently large portfolios there is no scope for further risk reduction through active portfolio management. However, if the firm parameters come from different distributions, say for different sectors or countries, then further risk reduction is possible, even asymptotically, by changing the portfolio weights. In either case, neglecting parameter heterogeneity can lead to *under*estimation of expected losses. But, 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. We examine the impact of sectoral and geographic diversification on credit losses empirically using returns for firms in the U.S. and Japan across seven sectors and find that ignoring this heterogeneity results in far riskier credit portfolios. Risk, is reduced significantly when parameter heterogeneity is properly taken into account.

*JEL Classifications*: C33, G13, G21.

*Key Words*: Risk management, correlated defaults, credit loss distributions, heterogeneity, diversification.

# 1 Introduction

The distinction between systematic and idiosyncratic risk is an integral part of the canon of corporate finance. The simple capital asset pricing model (CAPM) is perhaps its best known form. Idiosyncratic risk is readily diversified, leaving the investor exposed to systematic risk, the non-diversifiable component. But firms have different sensitivities to systematic risk, and systematic risk itself may be multi-dimensional with distinct risk types originating in specific industries, sectors or regions. In general, the potential for portfolio diversification then is driven broadly by these two characteristics: the degree to which systematic risk factors are correlated with each other and the degree of dependence of individual firms on different risk factors.

Although this paradigm has been developed for the analysis of risk in liquid market assets, it is nevertheless relevant to an investor in less liquid credit assets where obligor default is an event of particular interest. Models of the joint distribution of losses from a portfolio of credit assets form the cornerstone for a variety of applications in finance, from credit risk management to the pricing of credit assets such as CDOs (collateralized debt obligations) and credit derivatives. Credit risk analysis introduces another source of heterogeneity, namely the default threshold. This may vary across firms due, for instance, to different capital structures, and across countries because of, say, different bankruptcy laws.

In this paper we examine the scope for and limits of diversification for a credit portfolio. We derive analytical results for the portfolio loss distribution and present some complementary empirical findings. We do so for a wide set of different assumptions that underlie different aspects of the analysis of credit risk. The theoretical results indicate that if the firm parameters are heterogeneous but come from a common distribution, there is no scope for further risk reduction for a sufficiently large portfolio, i.e. one where idiosyncratic risk has already been diversified away. This would preclude gains from active portfolio management by changing the exposure weights (unless the portfolio is small, of course). However, if the firm parameters come from different distributions, say for different sectors or countries, then risk reduction is possible, even asymptotically, by changing the portfolio weights. In either case, neglecting parameter heterogeneity can lead to *under*estimation of expected losses (EL). But once EL is controlled for, neglecting parameter heterogeneity can lead to *over*estimation of unexpected losses or risk, whether measured by loss volatility, unexpected loss (UL), or value-at-risk (VaR).

Some of the analytical results are illustared using a large, two country (U.S. and Japan) portfolio. Return regressions with different degrees of parameter heterogeniety are estimated recursively using ten-year rolling estimation windows, with the loss distributions simulated for six *out-of-sample* one-year periods. The results are found to be robust across the six years. It is shown that, for a given EL, risk is significantly reduced when parameter heterogeneity is taken into account. Importantly, the introduction of parameter heterogeneity allows one to exploit diversification potential that

seems to exist in the selected sample portfolio. Since credit portfolio models impose conditional independence on firm returns, proper specification of the conditional mean (firm returns) and the default threshold work hand in glove. If conditional independence is violated, i.e. if there remains cross-sectional dependence in the residuals through poorly specified return regressions, then risk will be *under*estimated. Grouping firms by their credit rating, i.e. just seven groups instead of $N$ (the number of obligors), seems to capture most of the relevant heterogeneity needed for determining the default threshold. We also find that allowing for country specific factor loadings is important, but country specific factors seem to be less important by comparison.

The credit risk literature has recognized for some time the importance of modeling correlated or dependent defaults. Early treatment can be traced to the single (unobserved) factor model due to Vasicek (1987, 1991), which also forms the basis of New Basel Accord (BCBS, 2004) as outlined in detail by Gordy (2003). Extensions to multiple factors were proposed by Wilson (1997a,b) and Gupton, Finger and Bhatia (1997) in the form of the industry credit portfolio model CreditMetrics. For a summary of this and other industry models, see Saunders and Allen (2002), and for detailed comparisons, see Koyluoglu and Hickman (1998), Crouhy et al. (2000), and Gordy (2000). Practically all of these models are adaptations of Merton's (1974) options based approach, which develops a simple model of firm performance with a threshold value below which the firm defaults. In empirical applications the default threshold is modeled as a function of the firm's balance sheet. Not only is accounting information a noisy and possibly unreliable indicator of a firm's health, but in a multi-country setting it presents the additional challenges of different accounting standards and bankruptcy rules. In view of these measurement problems, Pesaran, Schuermann, Treutler and Weiner (2004), who link a credit risk model to a global macroeconometric multifactor model, propose an alternative estimation approach using firm-specific credit ratings and historical default frequencies.

A separate line of research has focused on correlated default intensities as in Schönbucher (1998), Duffie and Singleton (1999), Duffie and Gârleanu (2001) and Duffie and Wang (2004). There are a host of other approaches, including 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) and Frey and McNeil (2003).

Perhaps the most important distinction between our approach and the literature is around firm (or asset) heterogeneity: the risky asset pricing literature typically develops a model for a representative bond or firm.[1] Naturally, there will always be idiosyncratic or firm-specific differences,

---

[1] To be sure, one can find mention of multi-factor risk sensitivity (e.g. Duffie and Singleton (2003, Section 11.3.3)), but to our knowledge this topic has received at best casual treatment.

also allowed for in the risky asset pricing models. But our interest is in explicitly allowing for firm heterogeneity with respect to both the default threshold (or distance to default) and systematic risk sensitivity, an important dimension of diversification. Along the way we will be able to derive fat-tailed correlated losses from Gaussian (i.e. non-fat-tailed) risk factors and explore the potential for (and limits of) cross-sector and/or cross-country risk diversification.

The important technical problem of deriving the joint distribution of losses has received considerable attention of late. An early contribution was made by Vasicek in 1987, and then elaborated on in 1991, with the introduction of a single risk factor model with equicorrelated homogeneous exposures. Our model builds on Vasicek's results.

All of these studies in the first instance focus their modeling attention on the cumulative distribution function (CDF) of losses rather than its density. However, by working with CDFs, the researcher is typically limited to those with closed form solutions. Our strategy is to work instead with the probability density function (PDF) of losses allowing for broader generalizations, and these PDFs then may, or may not, have corresponding analytical CDFs. If not, this is not necessarily a problem as numerical solutions are often easily obtainable.

The plan for the remainder of the paper is as follows: Section 2 introduces the basic model of firm value and default, and Section 3 provides more detail regarding the specification and identification of the default threshold. Section 4 considers the problem of correlated defaults. Section 5 derives the portfolio loss distribution under different heterogeneity assumptions, starting with the simple case of a homogeneous portfolio as introduced by Vasicek. In Section 6 we derive asymptotic expressions for the loss density and distribution functions for normal and Student $t$ distributed firm innovations, again under both parameter homogeneity and heterogeneity assumptions. In Section 7 we allow loss-given-default to be correlated with the systematic risk factor(s), and the potential of sectoral and geographic diversification is discussed in Section 8. Section 9 explores the impact of heterogeneity empirically using returns for firms in the U.S. and Japan across seven sectors and analyzing the resulting loss distributions by simulation. Section 10 provides some concluding remarks.

## 2    Firm Value and Default

Much of the research on credit risk modelling 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.

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, $D_{i,t+h}$. A more nuanced approach is taken by the first-passage models (e.g. Black and Cox, 1976) where default would occur the first time that $V_{i,t}$ falls below a default boundary (or threshold) over the period $t$ to $t + h$.[2] The default probabilities are computed with respect to the probability distribution of asset values at the terminal date, $t + h$ in the case of the Merton model, and over the period from $t$ to $t + h$ in the case of the first-passage model. Therefore, the Merton approach may be thought of as a European option and the first-passage approach as an American option. Our analysis can be adapted to both of these models, but in what follows we focus on Merton's specification.

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

$$V_{i,t+h} \leq D_{i,t+h}.$$

Because default is costly and violations to the absolute priority rule in bankruptcy proceedings are common, in practice shareholders have an incentive to put the firm into receivership even before the equity value of the firm hits the zero value.[3] 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.[4] 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. This distinction is particularly relevant in the banking-borrower relationship we seek to characterize.[5]

As a result we shall assume that default takes place if

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

where $C_{i,t+h}$ is a positive default threshold which could vary over time and with the firm's particular characteristics (such as region or industry sector). Natural candidates include quantitative factors such as leverage, profitability, firm age (most of which appear in models of firm default), as well as more qualitative factors such as management quality.[6]

---

[2] For a review of these models, see, for example, Lando (2004, Chapter 3). More recent modeling approaches also allow for strategic default considerations, as in Mella-Barral and Perraudin (1997).

[3] 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.

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

[5] An excellent example of the joint borrower-lender decision process is given by Lawrence and Arshadi (1995).

[6] 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).

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

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

with a non-zero drift, $\mu_i$, and idiosyncratic Gaussian innovations with a zero mean and firm-specific volatility, $\sigma_{\xi_i}$. Consequently, the value of firm $i$ at time $t + h$ is

$$\ln(E_{i,t+h}) = \ln(E_{it}) + h\mu_i + \sum_{s=1}^{h} \xi_{i,t+s}.$$

and by (2) default occurs if

$$\ln(E_{i,t+h}) = \ln(E_{i,t}) + h\mu_i + \sum_{s=1}^{h} \xi_{i,t+s} < \ln(C_{i,t+h}), \tag{4}$$

or if the $h$-period change in equity value or return falls below the log-threshold-equity ratio, $\lambda_{i,t+h}$, defined by

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

Equation (5) tells us that the relative (rather than absolute) decline in firm value must be large enough over the horizon $h$ to result in default. Using (4), default occurs if

$$h\mu_i + \sum_{s=1}^{h} \xi_{i,t+s} < \lambda_{i,t+h},$$

Therefore, under (3) the probability that firm $i$ defaults at the terminal date $t + h$ is given by

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

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

# 3  Specification and Identification of Default Thresholds

Equation (6) provides a functional relationship between a firm's equity returns (as characterized by $\mu_i$ and $\sigma_{\xi_i}$), its default threshold, $\lambda_{i,t+h}$, and the default probability, $\pi_{i,t+h}$. In the case of publicly traded companies, $\mu_i$ and $\sigma_{\xi_i}$ can be consistently estimated from market returns based on historical data using either rolling or expanding observation windows. In general, however, $\lambda_{i,t+h}$ and $\pi_{i,t+h}$ can not be directly observed. One possibility would be to use balance sheet and other accounting data to estimate $\lambda_{i,t+h}$. This approach, for instance, is taken by KMV. But as argued in Pesaran, Schuermann, Treutler and Weiner (2004, PSTW), the accounting information are likely to be noisy and might not be all that reliable due to information asymmetries of managers, share-, and

debtholders.[7] Moreover, in a multi-country setting, the accounting based route presents additional challenges such as different accounting standards and bankruptcy rules that exist across countries. In addition to accounting data, other firm characteristics, such as leverage, firm age and perhaps size, and management quality could also be important in the determination of default thresholds that are quite difficult to observe. The accounting-based estimates of $\lambda_{i,t+h}$ are also likely to be highly data intensive, often requiring propriety information not readily available to general academic and professional researchers. 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.

Suppose that at the end of period $t$ firm $i$ is assigned a credit rating which we denote by $\mathcal{R}_t$. Typically $\mathcal{R}_t$ may take on values such as 'Aaa', 'Aa', 'Baa',..., 'Caa' in Moody's terminology, or 'AAA', 'AA', 'BBB',..., 'CCC' in Standard & Poor's (S&P) and Fitch's terminology. Suppose also that over the period, $t$ to $t + h$ the observed default frequency of $\mathcal{R}$-rated firms is given by $\hat{\pi}_{\mathcal{R},t+h}$. Therefore, under (3) and assuming that the number of $\mathcal{R}$-rated firms are sufficiently large, we have

$$\hat{\pi}_{\mathcal{R},t+h} = \sum_{i \in \mathcal{R}_t} w_{it} \Phi \left( \frac{\lambda_{i,t+h} - h \, \bar{\mu}_i}{\bar{\sigma}_{\xi_i} \sqrt{h}} \right), \tag{7}$$

where $\bar{\mu}_i$ and $\bar{\sigma}_{\xi_i}$ are the unconditional estimates of $\mu_i$ and $\sigma_{\xi_i}$ obtained using observations on firm-specific returns up to the end of period $t$, and $w_{it}$ is the weight of the $i^{th}$ firm in the portfolio of $\mathcal{R}$-rated firms at the end of period $t$, with $\sum_{i \in \mathcal{R}_t} w_{it} = 1$. The number of $\mathcal{R}$-rated firms at the end of period $t$ will be denoted by $N_{t\mathcal{R}}$.

The consistency of the above estimating equation requires $w_{it}$ to be pre-determined and non-dominating. Clearly, other grouping of firms can also be entertained. For example, firms can be grouped by industry or geographical regions as well as by their credit ratings. It would also be possible to consider averaging over firms with particular rating histories. For example, all firms that stayed within the same rating category for two successive periods (quarters, say), or those firms showing an improvement (deterioration) in their credit ratings over the two successive periods, $t - 1$ and $t$ could be included as separate categories (types). In considering these and many other "types" three important considerations ought to be born in mind. Firstly, the types should be reasonably homogeneous from the stand-point of default. Secondly, the number of firms of the same type must be sufficiently large so that the estimating equation (7) holds. Thirdly, there must be non-zero incidence of defaults across firms of the same type, namely $\hat{\pi}_{\mathcal{R},t+h} \neq 0$. Within type homogeneity is required since equation (7) contains $N_{t\mathcal{R}}$ unknown threshold parameters, $\lambda_{i,t+h}$ $i \in \mathcal{R}_t$. Their identification would require imposing a certain degree of homogeneity restrictions across the parameters, and/or one could find new moment conditions that relate the default thresholds to

---

[7]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.

the other characteristics of the empirical distribution of firm defaults. In what follows we consider two alternative exact identification schemes:

1. Within type homogeneity of defaults thresholds, namely

$$\lambda_{i,t+h} = \lambda_{\mathcal{R},t+h}, \quad \text{for all } i \in \mathcal{R}_t. \tag{8}$$

2. Within type homogeneity of distance-to-default

$$DD_{i,t+h} = \frac{\lambda_{i,t+h} - h\,\bar{\mu}_i}{\bar{\sigma}_{\xi_i}\sqrt{h}} = DD_{\mathcal{R},t+h}, \quad \text{for all } i \in \mathcal{R}_t. \tag{9}$$

Under the first identification scheme, the common default threshold, $\lambda_{\mathcal{R},t+h}$, can be obtained by solving the following non-linear equation in $\lambda_{\mathcal{R},t+h}$

$$\sum_{i \in \mathcal{R}_t} w_{it} \Phi\left(\frac{\lambda_{\mathcal{R},t+h} - h\,\bar{\mu}_i}{\bar{\sigma}_{\xi_i}\sqrt{h}}\right) - \hat{\pi}_{\mathcal{R},t+h} = 0 \tag{10}$$

It is easily seen that this equation has a unique, finite solution so long as $\hat{\pi}_{\mathcal{R},t+h} \neq 0$. Under the second identification scheme

$$\hat{\lambda}_{i,t+h} = DD_{\mathcal{R},t+h}\,\bar{\sigma}_{\xi_i}\sqrt{h} + h\,\bar{\mu}_i, \quad \text{for } i \in \mathcal{R}_t, \tag{11}$$

where

$$DD_{\mathcal{R},t+h} = \Phi^{-1}\left(\hat{\pi}_{\mathcal{R},t+h}\right), \tag{12}$$

and $\Phi^{-1}(\cdot)$ is the inverse of the cumulative distribution function of the standard normal.[8] Once again the estimated default thresholds, $\hat{\lambda}_{i,t+h}$, will be finite so long as $\hat{\pi}_{\mathcal{R},t+h} \neq 0$. Out 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 companies. 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 $DD$s 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 (11), with a brief discussion of the sensitivity of the results to other choices of $\lambda_{i,h+t}$. In the theoretical discussions that follows we shall assume that the firm-specific default thresholds are given, and do not consider the effects of their sampling uncertainty on the analysis of loss distributions.

---

[8]Note that $\Phi^{-1}(\pi_{i,t+h}) < 0$ for $\pi_{i,t+h} < 0.5$. In practice $\pi_{i,t+h}$ tends to be quite small.

# 4    Cross Firm Default Correlations

In the context of the Merton model the cross firm default correlations can be introduced by assuming the shocks to the firm asset values, $\xi_{i,t+1}$, defined by (3), to have the following multifactor structure

$$\xi_{i,t+1} = \boldsymbol{\gamma}_i' \mathbf{f}_{t+1} + \sigma_i \varepsilon_{i,t+1}, \quad \varepsilon_{i,t+1} \sim iid(0,1) \tag{13}$$

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 $i$. The common factors could be treated as unobserved or observed through macroeconomic variables such as output growth, inflation, interest rates and exchange rates.[9] In what follows we suppose the factors are unobserved, distributed independently of $\varepsilon_{i,t+1}$, and have constant variances.[10] 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 $m$.[11]

The above multifactor model plays a central role in the analysis of market risk, and its use in credit risk analysis seems a natural step towards a more cohesive understanding of the two types of risks and their relationships to one another. A homogeneous version of the factor model has also been used extensively for the analysis of credit portfolio risk by Vasicek (1987, 1991), as we shall see to good effect. But under homogeneity of factor loadings where $\boldsymbol{\gamma}_i = \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}_i' \mathbf{f}_{t+1} = \boldsymbol{\gamma}' \mathbf{f}_{t+1}$, then the distinction between a one factor and multifactor models will be redundant. Therefore, for multifactor analysis it is essential that we allow the factor loadings to be heterogeneous across firms.

Using (13) in (3) we now have

$$\ln(E_{i,t+1}) - \ln(E_{it}) = r_{i,t+1} = \mu_i + \boldsymbol{\gamma}_i' \mathbf{f}_{t+1} + \sigma_i \varepsilon_{i,t+1}, \tag{14}$$

and it is easily seen that

$$\sigma_{\xi_i}^2 = \boldsymbol{\gamma}_i' \boldsymbol{\gamma}_i + \sigma_i^2. \tag{15}$$

The presence of the common factors also introduces a varying degree of asset return correlations across firms, which in turn leads to 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, $\sigma_i$, the values of the default thresholds, $\lambda_{i,t+1}$, and the

---

[9] 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.

[10] 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 volatilities appear to be rather limited and of second order importance.

[11] The issues concerning the empirical implementation of the multifactor models in the context of credit risk models will be discussed in Section 9.

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

$$\rho_{ij} = \frac{\gamma_i' \gamma_j}{\left(\sigma_i^2 + \gamma_i' \gamma_i\right)^{1/2} \left(\sigma_j^2 + \gamma_j' \gamma_j\right)^{1/2}} = \frac{\delta_i' \delta_j}{\left(1 + \delta_i' \delta_i\right)^{1/2} \left(1 + \delta_j' \delta_j\right)^{1/2}}, \tag{16}$$

where $\delta_i = \gamma_i / \sigma_i$, the standardized factor loadings. Assuming $\delta_i$ and $\delta_j$ are independently distributed the average pair-wise correlation of asset returns is fixed and given by

$$E\left(\rho_{ij}\right) = E\left(\frac{\delta_i'}{\sqrt{1 + \delta_i' \delta_i}}\right) E\left(\frac{\delta_j}{\sqrt{1 + \delta_j' \delta_j}}\right).$$

The cross correlation of firm defaults, which we denote by $\rho_{ij}^*$, is more complicated to derive. Let $z_{i,t+1}$ to be the default outcome for firm $i$, over a single period such that[12]

$$z_{i,t+1} = I\left(\lambda_i - r_{i,t+1}\right), \tag{17}$$

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

$$\rho_{ij}^* = \frac{E\left(z_{i,t+1} z_{j,t+1}\right) - \pi_{i,t+1} \pi_{j,t+1}}{\sqrt{\pi_{i,t+1}(1 - \pi_{i,t+1})}\sqrt{\pi_{j,t+1}(1 - \pi_{j,t+1})}} \tag{18}$$

where $\pi_{i,t+1} = E\left(z_{i,t+1}\right)$ is firm $i's$ default probability over the period $t$ to $t + 1$. Using the return equation, (14), and the default criterion, (??), it is easily seen that $\rho_{ij}^* = 0$ if $\gamma_i = 0$ for all $i$ and $j$; namely conditional independence in returns carries over to defaults. For non-zero factor loadings relatively simple expressions for $\rho_{ij}^*$ can be obtained assuming $(\mathbf{f}_{t+1}, \varepsilon_{i,t+1})$ have a Gaussian distribution. In this case

$$\pi_{i,t+1} = \pi_i = \Phi\left(\frac{\lambda_i - \mu_i}{\sqrt{\sigma_i^2 + \gamma_i' \gamma_i}}\right). \tag{19}$$

The argument of $\Phi(\cdot)$ in (19) is commonly referred to as a "distance to default" $(DD)$ such that $DD_i = \Phi^{-1}(\pi_i)$. See also (9) and (12). To derive an expression for $E\left(z_{i,t+1} z_{j,t+1}\right)$ we first note that conditional on $\mathbf{f}_{t+1}$, $z_{i,t+1}$ and $z_{j,t+1}$ are independently distributed and

$$\begin{aligned} E\left(z_{i,t+1} z_{j,t+1}\right) &= E_f\left[E\left(z_{i,t+1} z_{j,t+1} \,|f_{t+1}\right)\right] \\ &= E_f\left[E\left(z_{i,t+1} \,|f_{t+1}\right) E\left(z_{j,t+1} \,|f_{t+1}\right)\right]. \end{aligned} \tag{20}$$

Also

$$\begin{aligned} E\left(z_{i,t+1} \,|\mathbf{f}_{t+1}\right) &= E\left(I\left(\lambda_i - \mu_i - \gamma_i' \mathbf{f}_{t+1} - \sigma_i \varepsilon_{i,t+1}\right) \,|\mathbf{f}_{t+1}\right) \\ &= \Phi\left(\frac{\lambda_i - \mu_i - \gamma_i' \mathbf{f}_{t+1}}{\sigma_i}\right) = \Phi\left(a_i - \delta_i' \mathbf{f}_{t+1}\right), \end{aligned}$$

---

[12]To simplify the exposition, and without any loss of generality, we set $h = 1$ and assume that default thresholds are time-invariant. These assumptions can be readily relaxed.

where as before $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i/\sigma_i$ and $a_i = \sigma_i^{-1}(\lambda_i - \mu_i)$.[13] Hence, unconditionally

$$E\left(z_{i,t+1}z_{j,t+1}\right) = E_f\left[\Phi\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1}\right)\Phi\left(a_j - \boldsymbol{\delta}_j'\mathbf{f}_{t+1}\right)\right], \tag{21}$$

where the expectations are now taken with respect to the distribution of the common factors, $\mathbf{f}_{t+1}$.

## 5  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]

$$\sum_{i=1}^{N} w_{it} = 1, \ \sum_{i=1}^{N} w_{it}^2 = O\left(N^{-1}\right), \ w_{it} \geq 0. \tag{22}$$

A sufficient condition for (22) 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] In what follows, for expositional simplicity, we shall be suppressing the time subscript on the portfolio weights unless they are specifically needed. It is straightforward to allow the weights to be time varying, given that at any point in time they are typically predetermined.

Suppose further that loss-given-default (LGD) of obligor $i$ is denoted by $\varphi_{i,t+1}$ which lies in the range $[0, 1]$.[16,17] Under this set-up the portfolio loss over the period $t$ to $t+1$ is given by

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it}\varphi_{i,t+1}z_{i,t+1}. \tag{23}$$

In cases where for each $i$, $\varphi_{i,t+1}$ and $z_{i,t+1}$ are independently distributed, the analysis can be conducted conditional on given values of LGD. In such a case the $\varphi_{i,t+1}$'s could be treated as fixed values and absorbed in the portfolio weights without loss of generality. However, a more interesting, and arguably practically more relevant case, arises where $\varphi_{i,t+1}$ and $z_{i,t+1}$ are correlated through common business cycle effects. This case presents new technical difficulties and is addressed briefly in Section 7. Until then, and without loss of generality, let $\varphi_{i,t+1} = 1 \ \forall i, t$, meaning that a defaulted

---

[13]Note that $a_i$ reduces to the distance to default, $DD_i$, defined above when $\gamma_i = \mathbf{0}$.

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

[15]Conditions (22) 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. See BCBS (2001, Ch.8).

[16]LGD is often modelled by assuming that $\varphi_{i,t+1}$ follows a Beta distribution across $i$ with parameters calibrated to match the mean and standard deviation of historical observations on the severity of credit losses.

[17]Lenders, be they banks or bondholders, often experience a technical default (e.g. a missed coupon payment or a breach of a covenant) without a loss, resulting in LGD = 0. Moreover, it is possible for LGD < 0, meaning that recovery exceeds exposure (i.e. > 100%). This may arise when the coupon rate of the debt instrument (loan or bond) exceeds the current interest rate.

asset has no recovery value, and write (23) as

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it} z_{i,t+1}. \tag{24}$$

The probability distribution function of $\ell_{N,t+1}$ can now be derived both conditional on an information set, $\mathcal{I}_t$, available at time $t$, or unconditionally. The two types of distributions coincide when the factors, $\mathbf{f}_{t+1}$, are assumed to be serially uncorrelated, a case often maintained in the literature. In this paper we consider 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 (proxies) when they are unobserved. A simple example of a dynamic factor model is the Gaussian vector autoregressive specification

$$\mathbf{f}_{t+1} = \mathbf{\Lambda}\mathbf{f}_t + \boldsymbol{\eta}_t, \ \boldsymbol{\eta}_t \mid \mathcal{I}_t \sim iidN(\mathbf{0}, \mathbf{\Omega}_{\eta\eta}), \tag{25}$$

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

$$Var\left(\mathbf{f}_{t+1} \mid \mathcal{I}_t\right) = \sum_{s=0}^{\infty} \mathbf{\Lambda}^s \mathbf{\Omega}_{\eta\eta} \mathbf{\Lambda}'^s = \mathbf{I}_m. \tag{26}$$

Along with much of the literature on credit risk, the focus of our analysis will be on the limit distribution of $\ell_{N,t+1} \mid \mathcal{I}_t$, as $N \to \infty$. The limit properties of this conditional loss distribution establishes the degree to which diversification of the credit portfolio is possible.[18] Not surprisingly, the limit distribution of the credit portfolio will depend on the nature of the return process $\{r_{i,t+1}\}$ and the extent to which the returns are cross-sectionally correlated.

## 5.1 Credit Risk under Firm Homogeneity

Vasicek (1987) was among 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, $\lambda_i = \lambda$, the same firm-specific volatilities, $\sigma_i = \sigma$, and zero unconditional returns, $\mu_i = 0$. As noted earlier a multifactor model with homogeneous factor loadings is equivalent to a single factor model. Under Vasicek's homogeneity assumptions we have

$$r_{i,t+1} = \gamma f_{t+1} + \sigma \, \varepsilon_{i,t+1},$$

where the single factor $f_{t+1}$ is also assumed to be serially uncorrelated. In this model the pair-wise asset return correlations, $\rho_{ij}$, will be identical for all obligor pairs in the portfolio and is given by

$$\rho_{ij} = \rho = \frac{\gamma^2}{\sigma^2 + \gamma^2}. \tag{27}$$

---

[18]The concept of "diversity" of financial markets has been recently discussed by Fernholz, Karatzas and Kardaras (2003), who provide a formal analysis in the context of the standard geometric Brownian motion model of asset returns.

Furthermore, since default depends on the sign of $\lambda - r_{i,t+1} = \lambda - (\gamma f_{t+1} + \sigma \, \varepsilon_{i,t+1})$, and not its magnitude, without loss of generality the normalization, $\sigma^2 + \gamma^2 = 1$ is often used in the literature, thus yielding $\gamma = \pm\sqrt{\rho}$. The remaining parameter, $\lambda$, is then calibrated to a pre-specified default probability, $\pi$, assuming a joint Gaussian distribution for $f_{t+1}$ and $\varepsilon_{i,t+1}$:

$$\begin{pmatrix} \varepsilon_{i,t+1} \\ f_{t+1} \end{pmatrix} \mid \mathcal{I}_t \sim iidN\left(\mathbf{0}, \mathbf{I}_2\right). \tag{28}$$

Under the above assumptions it is easily seen that

$$\pi = E\left(\ell_{N,t+1}\right) = \sum_{i=1}^{N} w_{it} E(z_{i,t+1}) = E(z_{i,t+1}) = \Pr\left(r_{i,t+1} \leq \lambda\right) = \Phi\left(\lambda\right).$$

Vasicek's model, therefore, takes the following simple form

$$r_{i,t+1} = \sqrt{\rho} f_{t+1} + \sqrt{1 - \rho} \, \varepsilon_{i,t+1}, \tag{29}$$

with the default threshold given by

$$\lambda = \Phi^{-1}\left(\pi\right), \tag{30}$$

so that the distance to default and default thresholds are the same. In Vasicek's model the pair-wise correlation of firm defaults, $\rho_{ij}^*$, is the same across all firms and is given by (see (18) and (21))

$$\rho_{ij}^* = \rho^*\left(\pi, \rho\right) = \frac{E\left\{\left[\Phi\left(\frac{\Phi^{-1}(\pi)}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}} f_{t+1}\right)\right]^2\right\} - \pi^2}{\pi(1 - \pi)}, \tag{31}$$

where expectations are taken with respect to the distribution of $f_{t+1}$, assumed to be $N(0,1)$. For example, for the standard parameter values $\pi = 0.01$, and $\rho = 0.30$, we have $\rho^* = 0.05$. In Figure 1, the top left chart labeled "Gaussian" (we shall return to the other charts in Section 5.4 below) provides simulated plots of $\rho^*\left(\rho, \pi\right)$ against $\rho$, for a few selected values of $\pi$. It is clear that the default correlation, $\rho^*$, is related non-linearly to $\rho$, and tends to be considerably lower than $\rho$. Also there is a clear tendency for the $(\rho^*, \rho)$ relationship to shift downwards as $\pi$ is reduced. For very small values of $\pi$, sizable default correlations are predicted by the double-Gaussian Vasicek model only for very high values of return correlations.[19]

## 5.2 Limits to Diversification - Vasicek's Model

Since the underlying returns are correlated, there is a non-zero lower bound to the unconditional loss variance, $Var\left(\ell_{N,t+1}\right)$, and full diversification will not be possible. Under the Vasicek model

$$Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \pi(1 - \pi)\left(\sum_{j=1}^{N} w_{jt}^2\right) + \pi(1 - \pi)\rho^*\left(\sum_{j \neq j'}^{N} w_{jt} w_{j't}\right),$$

---

[19] Determinants of $\rho^*$ in the case where the errors have Student-$t$ distribution with the same degree of freedom is discussed below. In particular, see (36).

where $\pi = E(z_{j,t+1})$ and $\rho^*$ is defined by (31) or by (36) below for Student-t errors. Since, $\sum_{j=1}^{N} w_j = 1$, it is easily seen that

$$\sum_{j=1}^{N} w_{jt}^2 + \sum_{j \neq j'}^{N} w_{jt} w_{j't} = 1,$$

so that

$$Var(\ell_{N,t+1} \mid \mathcal{I}_t) = \pi(1-\pi) \left\{ \rho^* + (1-\rho^*) \sum_{j=1}^{N} w_{jt}^2 \right\}. \tag{32}$$

Under the granularity condition, (22), for $N$ sufficiently large the second term in brackets become negligible, and $Var(\ell_{N,t+1})$ converges to the first term which will be non-zero for $\rho^* \neq 0$. Hence, in the limit, (32) converges to

$$\lim_{N \to \infty} Var(\ell_{N,t+1} \mid \mathcal{I}_t) = \pi(1-\pi)\rho^*. \tag{33}$$

For a finite value of $N$, the unexpected loss is minimized by adopting an equal weighted portfolio, with $w_j = 1/N$. For sufficiently large $N$, only the granularity condition (22) matters, and nothing can be gained by further optimization with respect of the weights, $w_j$.

## 5.3   Vasicek's Limit Distribution

The loss distribution for the perfectly homogeneous model is derived in Vasicek (1991, 2002) and Gordy (2000). Denoting the fraction of the portfolio lost to defaults by $x$, he obtains the following limiting density (as $N \to \infty$)

$$f_\ell(x \mid \mathcal{I}_t) = \sqrt{\frac{1-\rho}{\rho}} \left\{ \frac{\phi\left[ \frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(\pi)}{\sqrt{\rho}} \right]}{\phi\left[ \Phi^{-1}(x) \right]} \right\}, \text{ for } 0 < x \leq 1, \ \rho \neq 0, \tag{34}$$

where $\phi(\cdot)$ is the density function of a standard normal. The associated cumulative loss distribution function is

$$F_\ell(x \mid \mathcal{I}_t) = \Phi\left[ \frac{\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(\pi)}{\sqrt{\rho}} \right]. \tag{35}$$

As can be seen, Vasicek's limiting (as $N \to \infty$) credit loss distribution is fully determined by two parameters, namely the default probability, $\pi$, and the pair-wise return correlation coefficient, $\rho$. The former sets the expected loss of the portfolio, whilst the latter controls the shape of the loss distribution. In effect one parameter, $\rho$, controls all aspects of the loss distribution: its volatility, skewness and kurtosis. It would not be possible to calibrate two Vasicek loss distributions with the same expected and unexpected losses, but with different degrees of fat-tailedness, for example. Also, Vasicek's distribution does not depend on the portfolio weights so long as (22) is satisfied.

## 5.4 Default Correlations of Vasicek's Model under Non-Gaussian Distributions

It is well known that asset return distributions are fat-tailed and its neglect might result in under estimation of default correlations. In the context of Vasicek's model the importance of this issue can be investigated by considering $t$ distributions for the innovations ($\varepsilon_{i,t+1}$ and/or $f_{t+1}$) with low degrees of freedom, $t_v$, where $v > 2$ denotes the degrees of freedom of the distribution. When $\varepsilon_{i,t+1}$ is Gaussian but $f_{t+1} \sim iid\ t_v$, the computation of the default correlation coefficient, $\rho^*$, is straightforward and can be carried out using (31) with $f_{t+1}$ generated as draws from $iid\ t_v$.

However, the derivations are more complicated when $\varepsilon_{i,t+1}$ is $t$ distributed. In this case we must assume that $\varepsilon_{i,t+1}$ and $f_{t+1}$ are both $t$ distributed with the same degrees of freedom, otherwise $r_{i,t+1}$, given by (29), will have a non-standard distribution and the threshold parameter, $\lambda$, can not be derived analytically in terms of $\pi$. But when $\varepsilon_{i,t+1}$ and $f_{t+1}$ are both $t$ distributed with the same degrees of freedom, $v$, then $r_{i,t+1}$ will also be $t$ distributed with $v$ degrees of freedom and we have

$$\pi = \Pr\left(r_{i,t+1} \leq \lambda\right) = T_v\left(\lambda\right),$$

where $T_v\left(\cdot\right)$ denotes the cumulative distribution function of $t_v$, and hence, $\lambda = T_v^{-1}(\pi)$. Also

$$
\begin{aligned}
E\left(z_{i,t+1}\left|f_{t+1}\right.\right) &= E\left[I\left(\lambda - \sqrt{\rho}f_{t+1} - \sqrt{1-\rho}\,\varepsilon_{i,t+1}\right)\left|f_{t+1}\right.\right] \\
&= T_v\left(\frac{\lambda}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}}f_{t+1}\right).
\end{aligned}
$$

Using this result in (20) and then in (18) now yields

$$\rho^*\left(\pi,\rho,v\right) = \frac{E_f\left\{\left[T_v\left(\frac{T_v^{-1}(\pi)}{\sqrt{1-\rho}} - \sqrt{\frac{\rho}{1-\rho}}f_{t+1}\right)\right]^2\right\} - \pi^2}{\pi(1-\pi)}, \tag{36}$$

which is comparable to (31) obtained for Gaussian innovations. Expectations here are taken with respect to the distribution of $f_{t+1}$ assumed to be distributed as $t_v$.

Figure 1 contains simulated plots of $\rho^*\left(\rho,\pi,v\right)$ against $\rho$, for a few selected values of $\pi$ and for three values of $v$: 10, 5 and 3. As the innovations become increasingly fat-tailed, i.e. as $v$ declines, the curve becomes steeper meaning that default correlation $\rho^*$ increases more dramatically as return correlation, $\rho$, goes up. Moreover, differences in the default probability, $\pi$, matter less as the lines collapse on top of one another. Note the Gaussian case in the upper left representing $v = \infty$. Taken together it is clear that as innovations become more fat-tailed, the return correlation becomes the more important determinant of credit risk compared to the average default probability $\pi$, and they can potentially generate extremely large tail losses. For example, using (33) and (31), the unexpected loss of a Gaussian portfolio with $\pi = 0.01$, $\rho = 0.3$ is 0.021, whilst the unexpected loss of the same portfolio but with $t_3$ distributed shocks is 0.038.[20]

---

[20]This latter result is obtained using (33) and (36),

## 5.5 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's can be generalized. An empirical evaluation of the importance of allowing for firm heterogeneity in credit risk analysis is discussed in Section 9.

Under the heterogeneous multifactor return process, (14), the portfolio loss, $\ell_{N,t+1}$, can be written as

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it} I \left( a_i - \boldsymbol{\delta}_i' \mathbf{f}_{t+1} - \varepsilon_{i,t+1} \right), \tag{37}$$

where, as before $\boldsymbol{\delta}_i = \boldsymbol{\gamma}_i / \sigma_i$ are the standardized factor loadings, and $a_i = (\lambda_i - \mu_i)/\sigma_i$. In addition to allowing for parameter heterogeneity, we also relax the assumption that the common factors, $\mathbf{f}_{t+1}$, and the idiosyncratic shocks, $\varepsilon_{i,t+1}$, are normally distributed. Accordingly we assume that

$$\varepsilon_{i,t+1} \quad | \quad \mathcal{I}_t \sim iid\,(0,1), \text{ for all } i \text{ and } t,$$

$$\mathbf{f}_{t+1} \quad | \quad \mathcal{I}_t \sim iid\,(\boldsymbol{\mu}_t, \mathbf{I}_m), \text{ for all } t,$$

where under the dynamic factor model, (25), $\boldsymbol{\mu}_t = \Lambda \mathbf{f}_t$. Allowing $\boldsymbol{\mu}_t$ 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 $\boldsymbol{\mu}_t$ is usually set to zero.[21] 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 adopt the random coefficient model[22]

$$\boldsymbol{\theta}_i = \boldsymbol{\theta} + \mathbf{v}_i, \ \ \mathbf{v}_i \sim iid\,(\mathbf{0}, \Omega_{vv}), \text{ for } i = 1, 2, ..., N, \tag{38}$$

where

$$\boldsymbol{\theta}_i = \left( a_i, \boldsymbol{\delta}_i' \right)', \ \ \boldsymbol{\theta} = \left( a, \boldsymbol{\delta}' \right)', \ \ \mathbf{v}_i = \left( v_{ia}, \mathbf{v}_{i\delta}' \right)', \tag{39}$$

and

$$\boldsymbol{\Omega}_{vv} = \begin{pmatrix} \omega_{aa} & \boldsymbol{\omega}_{a\delta} \\ \boldsymbol{\omega}_{\delta a} & \boldsymbol{\Omega}_{\delta\delta} \end{pmatrix}, \tag{40}$$

is a semi-positive definite symmetric matrix, and $\mathbf{v}_i's$ are distributed independently of $(\varepsilon_{jt}, \mathbf{f}_t)$ 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

---

[21] With the possible exception of Wilson (1997a,b).

[22] Using (19), also note that

$$a_i = \frac{\lambda_i - \mu_i}{\sigma_i} = \sqrt{1 + \boldsymbol{\delta}_i' \boldsymbol{\delta}_i} \left[ \Phi^{-1}\left( \hat{\pi}_{\mathcal{R}} \right) \right], \text{ if } i \in \mathcal{R}.$$

But, to simplify the exposition we do not take account of the explicit dependence of $a_i$ on $\boldsymbol{\delta}_i$.

only up to their distributional properties described in (38). 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 to obtain estimates of $\boldsymbol{\theta}_i$, in which case $\boldsymbol{\theta}$ may need to be substituted for all borrowers $i$.

## 5.6 Limits to Unexpected Loss under Parameter Heterogeneity

The extent to which credit losses are diversifiable can be measured by $Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$, and can be obtained noting that in general

$$Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = E_f\left[Var\left(\ell_{N,t+1} \mid f_{t+1}, \mathcal{I}_t\right)\right] + Var_f\left[E\left(\ell_{N,t+1} \mid f_{t+1}, \mathcal{I}_t\right)\right]. \tag{41}$$

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 with 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. In particular, we have

$$E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = \sum_{i=1}^{N} w_{it} E\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right), \tag{42}$$

with the unexpected loss given by

$$Var\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = \sum_{i=1}^{N} w_{it}^2 Var\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right), \tag{43}$$

where the expected conditional default probability is

$$
\begin{aligned}
E\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) &= \Pr\left(\varepsilon_{i,t+1} \leq a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) \\
&= \int_{-\infty}^{\alpha_i - \beta_i f_{t+1}} f_\varepsilon(v)dv \equiv F_\varepsilon\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1}\right),
\end{aligned} \tag{44}
$$

and the conditional variance is

$$
\begin{aligned}
Var\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) &= E\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) - \left[E\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right]^2 \\
&= F_\varepsilon\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1}\right) - F_\varepsilon^2\left(a_i - \boldsymbol{\delta}_i'\mathbf{f}_{t+1}\right) \leq \frac{1}{4}.
\end{aligned} \tag{45}
$$

To integrate out the effects of the common factors on the variance of $\ell_{N,t+1}$, we shall make use of (41), (43) and (45) and note that

$$Var\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = \sum_{i=1}^{N} w_{it}^2 Var\left(z_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) \leq \frac{1}{4}\left(\sum_{i=1}^{N} w_{it}^2\right).$$

17

Hence, under (22)

$$E\left[Var\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right] \leq \frac{1}{4}\left(\sum_{i=1}^{N} w_{it}^2\right) \to 0, \text{ as } N \to \infty, \tag{46}$$

and in the limit the unexpected loss, $Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$, will be dominated by the second term in (41). Namely, we have

$$\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \lim_{N\to\infty} \left\{Var\left[E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right)\right]\right\}, \tag{47}$$

which is similar to Proposition 2 in Gordy (2003). This result clearly shows that when the portfolio weights satisfy the granularity condition, (22), the limit behavior of the unexpected loss does not depend on the weights $w_{it}$. Furthermore, this result holds irrespective of whether $a_i$ and $\boldsymbol{\delta}_i$ are homogeneous, or vary across $i$.

Under the random coefficient model, (38), the variance expression (47) can be obtained by integrating out the heterogeneous effects of $\alpha_i$ 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 (38) can be written as

$$\ell_{N,t+1} = \sum_{i=1}^{N} w_{it} I\left(a - \boldsymbol{\delta}' \mathbf{f}_{t+1} - \zeta_{i,t+1}\right), \tag{48}$$

where

$$\zeta_{i,t+1} = \varepsilon_{i,t+1} - \mathbf{v}_i' \mathbf{h}_{t+1} \tag{49}$$

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

$$\omega_{t+1}^2 = 1 + \mathbf{h}_{t+1}' \Omega_{vv} \mathbf{h}_{t+1} \tag{50}$$

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

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

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

$$E\left(\ell_{N,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = F_{\varkappa}\left(\frac{\boldsymbol{\theta}' \mathbf{h}_{t+1}}{\omega_{t+1}}\right), \tag{51}$$

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

$$\varkappa_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t = \frac{\zeta_{i,t+1}}{\omega_{t+1}} \sim iid(0,1). \tag{52}$$

Therefore, using (47), we have[23]

$$\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = Var\left[F_{\varkappa}\left(\frac{\boldsymbol{\theta}'\mathbf{h}_{t+1}}{\omega_{t+1}}\right) \mid \mathcal{I}_t\right], \tag{53}$$

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 by (33).

As in the homogeneous case, it is also clear that the limit of $Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$ as $N \to \infty$ vanishes *if and only if* $\mathbf{f}_{t+1}$ conditional on $\mathcal{I}_t$ is non-stochastic. Restated, allowing the portfolio to grow without bound, i.e. $N \to \infty$, eliminates idiosyncratic but not systematic risk. In general, when the returns are cross-sectionally correlated, full diversification is not possible and $\ell_{N,t+1}$ converges to a random variable with a non-degenerate probability distribution.

The implication for credit risk management is clear: *changing the exposure weights that satisfy (22) 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 in Section 8 below.

# 6   Limit Behavior of Credit Loss Distribution

## 6.1   Homogeneous Case with Non-Gaussian Innovations

In order to show how our approach relates to that of Vasicek, we first consider the homogeneous parameter case but do not require $\mathbf{f}_{t+1}$ 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 $\boldsymbol{\delta}_i$ and $\mathbf{f}_{t+1}$ and denote them by $\delta$ 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 $\ell_{N,t+1} \mid \mathcal{I}_t$ 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 $f_{\varepsilon}(\cdot)$.

---

[23]Numerical values of $\lim_{N\to\infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right)$ can be obtained by stochastic simulations, for a given distribution of $\varkappa_{i,t+1}$.

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

$$\ell_{t+1} = F_\varepsilon (a - \delta f_{t+1}) . \tag{54}$$

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

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

where $f_{t+1}$ is given in terms of $\ell_{t+1}$, via (54), namely

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

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

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

Hence

$$f_\ell (\ell_{t+1} \mid \mathcal{I}_t) = \frac{f_f \left( \frac{a - F_\varepsilon^{-1}(\ell,t+1)}{\delta} \mid \mathcal{I}_t \right)}{|\delta| f_\varepsilon \left[ F_\varepsilon^{-1} (\ell_{t+1}) \right]} , \quad \text{for } 0 < \ell_{t+1} \leq 1. \tag{55}$$

### 6.1.1 Relation to Vasicek's Loss Distribution

The above results provide a simple generalization of Vasicek's one-factor loss density distribution given by (34), and reduces to it when $\mu_t = 0$, and assuming that the innovations, $f_{t+1}$ and $\varepsilon_{i,t+1}$ are both Gaussian. In this case

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

and

$$f_\ell (x \mid \mathcal{I}_t) = \frac{1}{|\delta|} \left\{ \frac{\phi \left[ \frac{a - \Phi^{-1}(x)}{\delta} \right]}{\phi \left[ \Phi^{-1}(x) \right]} \right\} , \quad \text{for } 0 < x \leq 1, \ |\delta| \neq 0 \tag{56}$$

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

$$\delta = \sqrt{\frac{\rho}{1 - \rho}}, \quad \text{for } \rho > 0, \tag{57}$$

and

$$\pi = \Phi \left( \frac{a}{\sqrt{1 + \delta^2}} \right) . \tag{58}$$

Hence

$$a = \frac{\Phi^{-1}(\pi)}{\sqrt{1 - \rho}}. \tag{59}$$

Using (57) and (59) in (56) now yields Vasicek's loss density given by (34) (note that $\phi(x) = \phi(-x)$).

Under the double-Gaussian assumption, the distribution of $\delta f_{t+1} + \varepsilon_{t+1}$ (conditional on $\mathcal{I}_t$) is also Gaussian and we have (with $\mu_t \neq 0$)

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

Using (59) and (57) the conditional mean loss can therefore be written as

$$E\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = \Phi\left[\Phi^{-1}\left(\pi\right) - \sqrt{\rho}\mu_t\right], \tag{60}$$

and reduces to $\pi$ only when $\mu_t = 0$. It is also interesting to note that under $\mu_t \neq 0$, Vasicek's loss density and distributions become

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \sqrt{\frac{1 - \rho}{\rho}}\left\{\frac{\phi\left[\sqrt{\frac{1-\rho}{\rho}}\Phi^{-1}(x) - \sqrt{\frac{1}{\rho}}\Phi^{-1}(\pi) + \mu_t\right]}{\phi\left[\Phi^{-1}(x)\right]}\right\}, \text{ for } 0 < x \leq 1, \ \rho > 0. \tag{61}$$

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

$$F_\ell\left(x \mid \mathcal{I}_t\right) = \Phi\left(\sqrt{\frac{1 - \rho}{\rho}}\Phi^{-1}(x) - \sqrt{\frac{1}{\rho}}\Phi^{-1}\left(\pi\right) + \mu_t\right). \tag{62}$$

Also

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

which shows that good news (a rise in $\mu_t$) increases the probability of loss falling below a given threshold, or equivalently reduces the probability of losses above a given thresholds, as to be expected.

### 6.1.2 Non-Gaussian Innovations

It is interesting to note that Vasicek and others in this literature first derive the loss distribution function and then obtain the density - whilst we obtain the density first and then integrate to obtain the distribution function. One of the advantages of our procedure (aside from lending itself readily to heterogeneous generalizations) is that it can be used to derive analytic loss densities for non-Gaussian idiosyncratic shocks as well as non-Gaussian conditional distributions of $f_{t+1}$. For example, in the case where idiosyncratic shocks are Gaussian but the conditional distribution of the common factor is $t$ distributed with $v$ degrees of freedom, we have

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \frac{t_v\left(\sqrt{\frac{v}{v-2}}\left[\frac{a - \Phi^{-1}(x)}{\delta} - \mu_t\right]\right)}{|\delta| \phi\left[\Phi^{-1}(x)\right]}, \text{ for } 0 < x \leq 1, \ |\delta| \neq 0, \ v > 2,$$

where

$$t_v\left(u\right) = \frac{1}{\sqrt{v}B(v/2, 1/2)}\left[1 + v^{-1}u^2\right]^{-(v+1)/2},$$

and $B(v/2, 1/2)$ is the beta function.[24]

Using (55), other loss distributions can also be generated for different choices of the probability densities of $f_{t+1}$ and $\varepsilon_{i,t+1}$, although they might not be analytically tractable. In such cases the loss distribution can be generated by stochastic simulations.

## 6.2   Loss Densities under Heterogeneous Parameters

In this section we examine the loss behavior when under firm parameter heterogeneity, introduced in Section 5.5. In this case, $\ell_{N,t+1}$, is given by (48): $\ell_{N,t+1} = \sum_{i=1}^{N} w_i I\left(a - \boldsymbol{\delta}'\mathbf{f}_{t+1} - \zeta_{i,t+1}\right)$. Since conditional on $\mathbf{f}_{t+1}$, the composite errors, $\zeta_{i,t+1} = \varepsilon_{i,t+1} - v_{ia} + \mathbf{v}'_{i\boldsymbol{\delta}}\mathbf{f}_{t+1}$, are independently distributed across $i$, then

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

where as before $F_{\varkappa}(\cdot)$ denotes the cumulative distribution function of the standardized composite errors, $\varkappa_{i,t+1}$, defined by (52), $\mathbf{h}_{t+1} = \left(1, -\mathbf{f}'_{t+1}\right)'$ and $\omega_{t+1}$ is given by (50). Once again the limiting distribution of credit loss depends on the conditional densities of $\zeta_{i,t+1}$ and $\mathbf{f}_{t+1}$. For example, if $\left(\varepsilon_{i,t+1}, v_{ia}, \mathbf{v}'_{i\boldsymbol{\delta}}\right)$ follows a multivariate Gaussian distribution, then $\varkappa_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t \sim iidN(0,1)$.

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, $\mathbf{f}$, by application of standard change-of-variable techniques to the non-linear transformation

$$x = F_{\varkappa}\left(\frac{a - \boldsymbol{\delta}'\mathbf{f}}{\sqrt{1 + \omega_{aa} - 2\boldsymbol{\omega}'_{\alpha\delta}\mathbf{f} + \mathbf{f}'\Omega_{\delta\delta}\mathbf{f}}}\right). \tag{63}$$

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 $\mathbf{f}$ is a scalar, $f$. Suppose $f = \psi(x)$ satisfies the transformation, (63), and note that

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

where

$$\left|\psi'(x)\right| = \frac{1}{|x'(f)|}.$$

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

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

---

[24] We verified that this loss density (and its cdf, computed numerically using quadrature methods) generates very similar values for the Gaussian in (56) for $v = 30$, and is nearly indistinguishable for $v = 60$.

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 $0 < x \le 1$ we have

$$f_\ell\left(x \mid \mathcal{I}_t\right) = \left|\frac{\left[1 + \omega_{aa} - 2\omega_{a\delta}\psi(x) + \omega_{\delta\delta}\psi^2(x)\right]^{3/2}}{\psi(x)\left(\delta\omega_{a\delta} - a\omega_{\delta\delta}\right) + a\omega_{a\delta} - \delta\left(1 + \omega_{aa}\right)}\right|\left\{\frac{\phi\left[\psi(x) - \mu_t\right]}{\phi\left[\Phi^{-1}(x)\right]}\right\}. \tag{64}$$

This generalizes the result obtained for the homogeneous case, (56), and reduces to it if we set $\omega_{aa} = \omega_{a\delta} = \omega_{\delta\delta} = 0$. It is also interesting to note that the above 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 (22) are satisfied.

## 6.3   Implications of Parameter Heterogeneity for the Loss Distribution

As the above results clearly show, parameter heterogeneity can significantly affect the expected and unexpected losses, as well as the whole shape of the loss distribution. An analysis of the effects of heterogeneity on loss distribution in the general case, however, is analytically complicated and is best carried out via stochastic simulations, an approach that we shall consider in Section 9 below. But useful insights can be gained by limiting the analysis to the effects of heterogeneity of the mean returns and/or default thresholds across firms, assuming the factor loadings and the error variances are the same across firms. In this case, using (37) and assuming $\gamma_i = \gamma$ and $\sigma_i = \sigma$, we have

$$\ell_{N,t+1} = I\left(a_i - \delta\ f_{t+1} - \varepsilon_{i,t+1}\right),$$

where $\delta = \gamma/\sigma$, and $a_i = (\lambda_i - \mu_i)/\sigma$. This set-up is sufficiently general to allow for possible heterogeneity in the mean returns, $\mu_i$, and/or default thresholds, $\lambda_i$. Suppose that $a_i$ follows the random coefficient model

$$a_i = a + v_i, \quad v_i \sim iid\ N(0, \sigma_v^2). \tag{65}$$

It is then easily seen that

$$E\left(\ell_{N,t+1}\right) = \pi = \sum_{i=1}^{N} w_{it}\Pr\left(\delta\ f_{t+1} + \varepsilon_{i,t+1} - v_i \le a\right) = \Phi\left(\frac{a}{\sqrt{1 + \delta^2 + \sigma_v^2}}\right), \tag{66}$$

and

$$\lim_{N\to\infty}\left[Var\left(\ell_{N,t+1}\right) \mid \mathcal{I}_t\right] = Var_f\left[\Phi\left(\frac{a - \delta\ f_{t+1}}{\sqrt{1 + \sigma_v^2}}\right) \mid \mathcal{I}_t\right]. \tag{67}$$

These results clearly show that both expected and unexpected losses are affected by mean return/threshold heterogeneity. In this relatively simple example the degree of heterogeneity is unambiguously measured by the size of $\sigma_v^2$ and it is easily seen that,

$$\frac{\partial \pi}{\partial \sigma_v^2} = \phi\left(\frac{a}{\sqrt{1 + \delta^2 + \sigma_v^2}}\right)\frac{-a/2}{\left(1 + \delta^2 + \sigma_v^2\right)^{3/2}}$$

is positive since in practice one would expect $a < 0$. Notice that the distance to default is

$$\frac{a}{\sqrt{1 + \delta^2 + \sigma_v^2}} = \Phi^{-1}(\pi), \tag{68}$$

and for values of $\pi$ relevant in credit risk management, $\Phi^{-1}(\pi) < 0$. Therefore, for typical values of $\pi$, *the effect of heterogeneity would be to increase expected losses.* The dependence of $\pi$ on $\sigma_v^2$ is monotonic, and the higher the degree of heterogeneity the larger will be $\pi$. This result also shows that neglecting heterogeneity in credit risk analysis can lead to under-estimation of mean losses and can have hard to predict consequences on higher moments, and in particular the tails of the loss distribution.

To examine the effect of heterogeneity on unexpected losses, we first control for the effect of changes in $\sigma_v^2$ on expected losses by setting $a = \sqrt{1 + \delta^2 + \sigma_v^2}\, \Phi^{-1}(\pi)$. From (66) it is clear that this choice of $a$ ensures that $E(\ell_{N,t+1}) = \pi$, irrespective of the value of $\sigma_v^2$. Using (67) it now follows that

$$\lim_{N \to \infty} [Var(\ell_{N,t+1}) \mid \mathcal{I}_t] = Var_f\left[\Phi\left(\Phi^{-1}(\pi)\sqrt{1 + \kappa^2} - \kappa f_{t+1}\right) \mid \mathcal{I}_t\right].$$

where

$$\kappa = \frac{\delta}{\sqrt{1 + \sigma_v^2}}.$$

Also, the pair-wise correlation coefficient, $\rho_{ij}$, in this case is given by

$$\rho_{ij} = \rho = \frac{\delta^2}{1 + \delta^2 + \sigma_v^2} = \frac{\kappa^2}{1 + \kappa^2}, \tag{69}$$

and as in the homogeneous case is the same across all $i$ and $j$. Hence, noting that $\kappa^2 = \rho/(1 - \rho)$, we have

$$\lim_{N \to \infty} [Var(\ell_{N,t+1}) \mid \mathcal{I}_t] = Var_f\left[\Phi\left(\frac{\Phi^{-1}(\pi)}{\sqrt{1 - \rho}} - \sqrt{\frac{\rho}{1 - \rho}} f_{t+1}\right) \mid \mathcal{I}_t\right].$$

Therefore, under $E(\ell_{N,t+1}) = \pi$, in the limit as $N \to \infty$ the unexpected loss depends on the degree of parameter heterogeneity, $\sigma_v^2$, only through the return correlation coefficient, $\rho$. From (69) note that for a given value of $\delta$ (the standardized threshold), $\rho$ is a decreasing function of $\sigma_v^2$. A rise in heterogeneity (or $\sigma_v^2$) reduces $\rho$, which in turn results in a reduction of unexpected losses. So, once expected losses are appropriately corrected to take account of the increased risk of dealing with a heterogeneous sample, that very heterogeneity can be exploited to achieve greater diversification of the credit risk portfolio. Indeed as we shall see in Section 9.4, simulation reveals that once expected losses are controlled for, ignoring parameter heterogeneity results in significant overestimation of credit risk, especially in the tails.

# 7 Allowing for Correlated Loss Severity

Consider now the more general case where the loss function is given by (23) and assume that LGD variables, $\varphi_{i,t+1}$, depend on the common factors $\mathbf{f}_{t+1}$ but conditional on $\mathbf{f}_{t+1}$,[25] $\varphi_{i,t+1}$ and $z_{i,t+1}$ are independently distributed.[26] A common formulation for LGD is the Beta distribution, convenient because it is bounded between 0 and 1, with two shape parameters that can be expressed in terms of the average and volatility of observed LGDs. Specifically, suppose that

$$f_\varphi\left(\varphi_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) \sim Beta(p_{t+1}, q_{t+1}),$$

where $Beta(p_{t+1}, q_{t+1})$ is a Beta distribution with parameters $p$ and $q$ that depend on $\mathbf{f}_{t+1}$, and possibly on $\mathcal{I}_t$. Note also that

$$E\left(\varphi_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = \frac{p_{t+1}}{p_{t+1} + q_{t+1}}.$$

We shall consider the following logistic form as an example of this mean function

$$E\left(\varphi_{i,t+1} \mid \mathbf{f}_{t+1}, \mathcal{I}_t\right) = \frac{e^{\boldsymbol{\kappa}' \mathbf{f}_{t+1}}}{1 + e^{\boldsymbol{\kappa}' \mathbf{f}_{t+1}}},$$

where $\boldsymbol{\kappa}$ measures the degree to which mean $LGD$ varies with the business cycle indicators, here represented by $\mathbf{f}_{t+1}$. Clearly other functional forms and business cycle indicators can also be used. Note that LGD is assumed not to vary systematically across exposures $i$ but only over $t$.[27]

Since $\left|\varphi_{i,t+1} z_{i,t+1}\right| \leq 1$, firstly the results (46) and (47) continue to hold. Also under the random coefficient model (38) we have

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

and hence

$$\lim_{N \to \infty} Var\left(\ell_{N,t+1} \mid \mathcal{I}_t\right) = Var\left[\left(\frac{e^{\boldsymbol{\kappa}' \mathbf{f}_{t+1}}}{1 + e^{\boldsymbol{\kappa}' \mathbf{f}_{t+1}}}\right) F_\varkappa\left(\frac{\boldsymbol{\theta}' \mathbf{h}_{t+1}}{\omega_{t+1}}\right) \mid \mathcal{I}_t\right].$$

Finally, the limit distribution in this case can also be derived noting that

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

---

[25] For empirical evidence of procyclical LGD, see Frye (2000), Altman et al. (2003) and Hu and Perraudin (2002).

[26] This does not seem to be a restrictive assumption.

[27] This formulation ties the two random variables of interest, default and $LGD$, to the same systematic risk factor(s) while allowing the factor loadings to be different. Again, the basic idea is that credit risk correlations/dependence, difficult if not impossible to observe directly, are modeled indirectly through the systematic factors. This is both conceptually desirable because we are using a structural model, as well as pragmatically and empirically useful since we focus the modeling effort where the data is dense and not sparse.

# 8 Possible Sectoral or Geographic Diversification

The results obtained so far provides the limits to risk diversification through inclusion of additional firms with different idiosyncratic characteristics. For the homogeneous case there is a lower bound to the unexpected loss given by

$$Var\left[F_\varepsilon\left(a - \delta f_{t+1}\right) \mid \mathcal{I}_t\right],$$

and for the heterogeneous case by

$$Var\left[F_\varkappa\left(\frac{\boldsymbol{\theta}'\mathbf{h}_{t+1}}{\omega_{t+1}}\right) \mid \mathcal{I}_t\right].$$

In both cases as $N \to \infty$, unexpected losses do not depend on the exposure weights, $w_i$. Furthermore, if $\mathbf{f}_{t+1}$ are serially independent (as is often assumed in the finance literature), then the above bounds hold unconditionally, namely the lower bound to risk diversification is given by

$$Var\left(\ell_{N,t+1}\right) > Var\left[F_\varkappa\left(\frac{\boldsymbol{\theta}'\mathbf{h}_{t+1}}{\omega_{t+1}}\right)\right].$$

Once idiosyncratic risk vanishes, there is no scope for active diversification through changes in the portfolio weights so long as $N$ is sufficiently large and $w_i$ satisfy the atomistic conditions, (22).

There might, however, be important possibilities for further diversification if we could group the firms into different categories with the parameters of each category having different distributions. One may think of these categories as different industries, sectors, or countries, for instance, whose sensitivities to the systematic risk factor are stochastic. As a simple example suppose there are $N = N_A + N_B$ firms grouped into country $A$ (say Japan) and country $B$ (say U.S.) such that

$$A \quad : \quad r_{Ai,t+1} = \mu_{Ai} + \boldsymbol{\gamma}'_{Ai}\mathbf{f}_{t+1} + \sigma_{Ai}\varepsilon_{Ai,t+1}, \;\; i = 1, 2, ..., N_A$$
$$B \quad : \quad r_{Bi,t+1} = \mu_{Bi} + \boldsymbol{\gamma}'_{Bi}\mathbf{f}_{t+1} + \sigma_{Bi}\varepsilon_{Bi,t+1}, \;\; i = 1, 2, ..., N_B,$$

where

$$\mu_{Ai} \;\; = \;\; \mu_A + v_{A\mu i}, \quad \mu_{Bi} = \mu_B + v_{B\mu i},$$
$$\boldsymbol{\gamma}_{Ai} \;\; = \;\; \boldsymbol{\gamma}_A + \mathbf{v}_{A\gamma i}, \quad \boldsymbol{\gamma}_{Bi} = \boldsymbol{\gamma}_B + \mathbf{v}_{B\gamma i}.$$

Thus, for example, fixed effects for Japanese firms $(A)$ are randomly distributed around a country mean, $\mu_A$, and the Japanese systematic factor loading is also randomly distributed around a country effect, $\boldsymbol{\gamma}_A$. Suppose further that those errors $\left(v_{A\mu i}, \mathbf{v}'_{A\gamma i}\right)'$ and $\left(v_{B\mu i}, \mathbf{v}'_{B\gamma i}\right)'$ are independently distributed:

$$\begin{pmatrix} v_{A\mu i} \\ \mathbf{v}_{A\gamma i} \end{pmatrix} \;\; \sim \;\; iid\left(\mathbf{0}, \boldsymbol{\Omega}^A_{vv}\right),$$

$$\begin{pmatrix} v_{B\mu i} \\ \mathbf{v}_{B\gamma i} \end{pmatrix} \;\; \sim \;\; iid\left(\mathbf{0}, \boldsymbol{\Omega}^B_{vv}\right).$$

Therefore cross-country or -sector dependence arises only through $\mathbf{f}_{t+1}$ and not through the parameter distributions themselves, although it is now possible that different factors could affect the firm returns in different countries or sectors.

Consider now the following credit portfolio composed of two separate portfolios each with weights $\varpi_t$ and $(1 - \varpi_t)$:

$$\ell_{N,t+1}^{(A,B)} = \varpi_t \sum_{i=1}^{N_A} w_{iA} I \left( \lambda_{iA} - r_{iA,t+1} \right) + (1 - \varpi_t) \sum_{i=1}^{N_B} w_{iB} I \left( \lambda_{iB} - r_{iB,t+1} \right),$$

where

$$\sum_{i=1}^{N_A} w_{iA} = \sum_{i=1}^{N_B} w_{iB} = 1,$$

$$\sum_{i=1}^{N_A} w_{iA}^2 \rightarrow 0, \sum_{i=1}^{N_B} w_{iB}^2 \rightarrow 0, \text{ as } N_A \text{ and } N_A \rightarrow \infty,$$

meaning the two sub-portfolios have a large number of relatively small exposures. We may compare the "joint" portfolio to the following "single-country" portfolios

$$\ell_{N_A,t+1}^{(A)} = \sum_{i=1}^{N_A} w_{iA} I \left( \lambda_{iA} - r_{iA,t+1} \right),$$

or

$$\ell_{N_B,t+1}^{(B)} = \sum_{i=1}^{N_B} w_{iB} I \left( \lambda_{iB} - r_{iB,t+1} \right).$$

It is now easily seen that the limit of the unexpected losses associated with these portfolios as $N_A, N_B \rightarrow \infty$, are given by

$$\lim_{N_A \rightarrow \infty} Var \left[ \ell_{N_A,t+1}^{(A)} \mid \mathcal{I}_t \right] = Var \left[ F_\varkappa \left( \frac{\boldsymbol{\theta}_A' \mathbf{h}_{t+1}}{\omega_{A,t+1}} \right) \mid \mathcal{I}_t \right] = V_{tA}, \tag{70}$$

$$\lim_{N_B \rightarrow \infty} Var \left[ \ell_{N_B,t+1}^{(B)} \mid \mathcal{I}_t \right] = Var \left[ F_\varkappa \left( \frac{\boldsymbol{\theta}_B' \mathbf{h}_{t+1}}{\omega_{B,t+1}} \right) \mid \mathcal{I}_t \right] = V_{tB},$$

$$\lim_{N_A,N_B \rightarrow \infty} Var \left[ \ell_{N,t+1}^{(A,B)} \mid \mathcal{I}_t \right] = \varpi_t^2 V_{tA} + (1 - \varpi_t)^2 V_{tB},$$

where $\boldsymbol{\theta}_A = (a_A, \boldsymbol{\delta}_A')'$, $\boldsymbol{\theta}_B = (a_B, \boldsymbol{\delta}_B')'$, and $\omega_{s,t+1}^2 = 1 + \mathbf{h}_{t+1}' \Omega_{vv}^s \mathbf{h}_{t+1}$, for $s = A, B$. Hence, the unexpected losses of the combined portfolio will be minimized with

$$\varpi_t^* = \frac{V_{tB}}{V_{tA} + V_{tB}}.$$

Note that both $N_A$ and $N_B$ individually need to be sufficiently large for idiosyncratic risk to vanish. Not surprisingly it is optimal to place a larger weight on the portfolio with a smaller unexpected loss conditional on $\mathcal{I}_t$. Using $\varpi_t^*$ we have

$$\lim_{N \rightarrow \infty} Var \left[ \ell_{N,t+1}^{(A,B)} \mid \mathcal{I}_t \right] = \frac{V_{tA} V_{tB}}{V_{tA} + V_{tB}},$$

which is smaller than either $V_{tA}$ or $V_{tB}$. Therefore, *the joint sectorally or geographically diversified portfolio will be less risky than either of the standalone portfolios A or B.*

# 9 An Empirical Application: Heterogeneity and Risk Diversification

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 public firms from the U.S. and Japan. We form these credit portfolios at the end of each year from 1997 to 2002 and then simulate portfolio losses for the following year. All of the simulation 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 time variation in the underlying parameters as well as the effects that such time variations might have on loss distributions.

## 9.1 Data and Portfolio Construction

The loss simulations require an estimate of the unconditional probability of default for each firm. These may be obtained at the level of the credit rating, $\mathcal{R}$, assigned to the firm by rating agencies such as Moody's, S&P or Fitch. Rating agencies effectively group firms into credit ratings such as $\mathcal{AAA}$, $\mathcal{BB}$, etc., by credit quality. In assessing the credit quality of a firm, a rating agency has access to, and presumably makes use of, information about a firm not necessarily available to outsiders. Given credit rating histories, one is able to compute default probabilities by rating, and with sufficient data, it is in principle possible to allow for more granularity and obtain estimates of $\pi$ by country and industry, by rating, but at present the available data do not allow for this. 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 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 six cohorts.

Credit ratings have an additional advantage in a multi-country context. They are designed to be directly comparable across countries, meaning that a $\mathcal{AA}$ in Japan is comparable to a $\mathcal{AA}$ in the U.S., thus controlling for heterogeneity in accounting standards and bankruptcy conventions across countries.

In order to be selected for inclusion in our portfolios, a firm needs a credit rating as well as 10 years of consecutive quarterly equity returns that match the rolling estimation window. The first estimation sample covers the ten years ending in December 31, 1997, the second ten-year period ends in December 31, 1998, and so on with the last estimation sample ending in December 31, 2002, providing us with six yearly recursive forecasts for the years 1998 to 2003, inclusive. For these recursions to be operational the initial portfolio requires firms to have a credit rating from

28

either Moody's or S&P at year-end 1997 as well as 40 consecutive quarters of equity returns, from 1988Q1 to 1997Q4. In case both ratings are available the S&P rating is chosen.[28] For the first sample or cohort (which ends in 1997) we have 211 Japanese firms and 628 U.S. firms; a portfolio of 839 firms in total. At the end of the following year the portfolio is rebalanced retaining surviving firms and augmented with additional new firms that have a rating at the end of that year, i.e. 1998, and also have 40 consecutive quarters of returns.[29] All returns are computed in U.S. dollar (USD). For Japanese firms this is done by subtracting the rate of change of yen/USD exchange rate from their yen-denominated returns.

To make the portfolio exposures representative of the rated universe in each country, we re-weight the portfolio exposures (in USD) by rating in the following manner. Suppose that each obligor begins with $100 of exposure. If 10% of all rated Japanese firms have a $\mathcal{BB}$ rating, but 15.6% of the Japanese firms in our portfolio are $\mathcal{BB}$-rated, then each of these firms will be given $63 ($\frac{10.0}{15.6} \cdot \$100$) of exposure. In the U.S. the difference in the ratings distribution across the two agencies is modest, but not so in Japan where Moody's rates more than twice as many firms as S&P. To address this issue we take the average of the two agencies' ratings distributions by rating for each country.[30]

The portfolio composition is adjusted annually starting with 1998 to reflect defaults, upgrades and downgrades which may have occurred during the year. Since migration matrices even at annual frequencies are diagonally dominant, with average staying probabilities exceeding 90% for investment grades, annual portfolio rebalancing seems a reasonable compromise between accuracy and computational burden; the alternative would be quarterly rebalancing. We also update the ratings distribution each year to allow for compositional changes in the universe of rated firms. For example, at the end of 1997 $\mathcal{B}$-rated firms made up only 1.54% of all rated Japanese firms, but by year-end 2002 this proportion had risen to 6.75%. Below in Table 1 we show the average ratings distribution for each country for 1997 and 2002. It becomes clear that there has been a systematic deterioration in average credit quality 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 1.23% for the year-end 1997 portfolio to 3.26% for the year-end 2002 portfolio.

---

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

[29] Our source of return data for U.S. firms is CRSP while for Japanese firms it is Datastream.

[30] The precise exposure allocation is as follows. Denote $FV_{ic}$ to be the (face value) exposure to firm $i$ in country $c$. The portfolio total nominal face value is $1bn. Then

$$FV_{ic} = \$1bn \cdot w_c \cdot \left(\frac{1}{N_c}\right) \cdot \theta(\mathcal{R})_c \ \text{ for } i \in \mathcal{R},$$

where $w_c$ is the share of the total portfolio for country $c$ (75% for the U.S., 25% for Japan), $N_c$ is the number of firms in country $c$, $\theta(\mathcal{R})_c$ is the rating representation adjustment. Note that $\theta(\mathcal{R})_c$ will vary across time to reflect compositional changes in the rated universe of firms in county $c$.

**Table 1: Ratings Distributions and Probabilities of Default**

| | 1997 | | | | 2002 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Ratings Distribution (%)** | | | | **Ratings Distribution (%)** | | | |
| **Credit Rating** | **Japan** | **U.S.** | **Both** | $\hat{\pi}_{\mathcal{R}}(\%)$ | **Japan** | **U.S.** | **Both** | $\hat{\pi}_{\mathcal{R}}(\%)$ |
| $\mathcal{AAA}$ | 4.80 | 2.86 | 3.35 | 0.001 | 0.75 | 2.37 | 1.97 | 0.001 |
| $\mathcal{AA}$ | 22.62 | 10.81 | 13.76 | 0.001 | 16.03 | 9.25 | 10.95 | 0.001 |
| $\mathcal{A}$ | 37.93 | 25.61 | 28.69 | 0.005 | 29.38 | 21.49 | 23.46 | 0.006 |
| $\mathcal{BBB}$ | 23.16 | 22.33 | 22.54 | 0.064 | 34.89 | 25.67 | 27.97 | 0.106 |
| $\mathcal{BB}$ | 9.94 | 16.30 | 14.71 | 0.481 | 11.85 | 15.72 | 14.75 | 0.630 |
| $\mathcal{B}$ | 1.54 | 19.79 | 15.22 | 3.343 | 6.75 | 19.62 | 16.40 | 5.429 |
| $\mathcal{CCC}$ | 0.00 | 2.31 | 1.73 | 36.49 | 0.35 | 5.89 | 4.50 | 49.78 |
| **Portfolio $\hat{\pi}$** | | | | **1.23** | | | | **3.26** |

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 by country at year-end 1997 are summarized below in Table 2.

**Table 2: Industry Breakdowns by Country**

(Percent of observations) at year-end 1997

| | U.S. | Japan |
|---|---|---|
| Agriculture, Mining & Construction | 5.3 | 8.5 |
| Communication, Electric & Gas | 16.7 | 6.2 |
| Durable Manufacturing | 22.1 | 34.1 |
| Finance, Insurance & Real Estate | 23.1 | 14.7 |
| Non-durable Manufacturing | 18.2 | 24.6 |
| Service | 4.8 | 6.6 |
| Wholesale & Retail Trade | 9.9 | 5.2 |
| **Total** | **100** | **100** |

## 9.2 Model Specifications

To explore the role of geographic and sectoral heterogeneity we introduce two new indices into the notation of the previous sections. Specifically, denote $r_{ijc,t+1}$ to be the return of firm $i$ in sector $j$ in country $c$ over the quarter $t$ to $t+1$, where $i = 1, ..., I_{jc}$, $j = 1, ..., I_c$, $c = 1, ..., C$. The

application will explore $C = 2$ countries (Japan and U.S.) and $I_c = 7$ sectors/industries in each country. Following (14), and relaxing the assumption that $\mathbf{f}_{t+1} \sim (\mathbf{0}, \mathbf{I}_m)$, made for expositional simplicity, the return equations relevant to our empirical applications can be written as

$$r_{ijc,t+1} = \alpha_{ijc} + \boldsymbol{\beta}'_{ijc}\mathbf{f}_{t+1} + u_{ijc,t+1}, \tag{71}$$

where $\mathbf{f}_{t+1} \sim (\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f)$, $\boldsymbol{\mu}_f$ 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 (14) we have

$$\mu_{ijc,t+1} = \alpha_{ijc} + \boldsymbol{\beta}'_{ijc}\boldsymbol{\mu}_f, \tag{72}$$

and

$$\xi_{ijc,t+1} = \boldsymbol{\beta}'_{ijc}(\mathbf{f}_{t+1} - \boldsymbol{\mu_f}) + u_{ijc,t+1}. \tag{73}$$

Note also that

$$\sigma^2_{\xi_{ijc}} = \boldsymbol{\beta}'_{ijc}\Sigma_f\boldsymbol{\beta}_{ijc} + \sigma^2_{ijc}, \tag{74}$$

where $\sigma^2_{ijc}$ is the variance of the idiosyncratic component, $u_{ijc,t+1}$.

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.[31] One obvious source of heterogeneity is geography or country. As we have two countries, we estimate each model specification first by pooling the U.S. and Japanese firms (referred to as the "pooled model" specification) and then by estimating a separate model by country (the "modeled separately" specification).

The empirical exercise involves a number of variations on the basic firm return equation given by (71) using market-cap weighted market returns for each country $\bar{r}_{c,t+1}$ as proxies for two of the possible $m$ common factors. Sector returns for a country $c$, $\bar{r}_{jc,t+1}$, $j = 1, ..., I_c$, are computed in a similar fashion, namely using the market-cap weighted average of firm returns in that sector.[32] The "global" market return index, $\bar{r}_{t+1}$, is made up of just the two countries U.S. and Japan and is simply the weighted sum of the two individual country returns,

$$\bar{r}_{t+1} = w_{US}\bar{r}_{US,t+1} + (1 - w_{US})\bar{r}_{JP,t+1}, \tag{75}$$

where $w_{US}$ measures the relative size of the U.S. economy. We estimate $w_{US}$ by taking the average U.S. share of PPP-denominated GDP over 1997-2002, and obtain $w_{US} = 0.75$. To obtain the global sector return $\bar{r}_{j,t+1}$ for a particular sector $j$ we proceed similarly to (75) and define

$$\bar{r}_{j,t+1} = w_{US}\bar{r}_{j,t+1,US} + (1 - w_{US})\bar{r}_{j,t+1,JP}. \tag{76}$$

---

[31] An application where the firm returns are linked to observable macroeconomic factors using a global VAR model is provided in PSTW.

[32] The weights for period $t + 1$ are based on the average of the market capitalization (in USD) at end of periods $t$ and $t + 1$.

The simplest model is the return specification assumed under the Vasicek model,

$$r_{ijc,t+1} = \alpha_c + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}, \tag{77}$$

with homogeneous error variances, $u_{ijc,t+1} \sim iidN(0, \sigma_c^2)$. For the pooled model, $\sigma_c^2 = \sigma^2, \alpha_c = \alpha$, $\beta_c = \beta$ and $\bar{r}_{c,t+1} = \bar{r}_{t+1}$ as in (75) for $c = US$ and $JP$.

Next is the fixed effect model where we allow $\alpha_{ijc}$ to vary across firms $i$ in each sector $j$ and country $c$. We examine two versions of the fixed effects model, one holding the error variances fixed across all firms $(\sigma_c^2)$ and the other allowing each firm to have its own error variance $(\sigma_{ijc}^2)$. When we hold the error variances fixed across all firms, the slope coefficient is unchanged from the one obtained under (77). When we allow each firm to have its own error variance, the slope coefficient is always less than the estimate obtained under (77).

The third model specification allows for full parameter heterogeneity where fixed effects, slopes and error variances are allowed to vary across firms. Here too we estimate two versions, where in the second we add an industry or sector factor so that each firm's returns is regressed on $\bar{r}_{c,t+1}$ as well as on $\bar{r}_{cj,t+1}$. To be clear, for the pooled model, $\bar{r}_{c,t+1} = \bar{r}_{t+1}$ as in (75), and $\bar{r}_{cj,t+1} = \bar{r}_{j,t+1}$ as in (76), for $c = US$ and $JP$.

The fourth and final model specification is the principal components (PCA) model. See Appendix A for further detailed account of this specification. Using the procedure outlined in Bai and Ng (2002), we extract $\hat{m}$ relevant principal components which capture most of the in-sample variation in firm returns. In the application, the procedure resulted in two factors for the U.S., three for Japan and three for the pooled model (i.e. when firm returns from the U.S. and Japan are pooled). 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 all eight model specifications that we consider.

**Table 3: Specifications of Return Equations and Default Thresholds for Separate Country Models,** $c = US, JP$

| Models | Descriptions | Returns | Default Thresholds |
|---|---|---|---|
| I | Vasicek | $r_{ijc,t+1} = \alpha_c + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_c / DD_c$ |
| II(a) | Fixed Effects $(\sigma^2)$ | $r_{ijc,t+1} = \alpha_{ijc} + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_c / DD_{c,i}$ |
| II(b) | Fixed Effects $(\sigma_i^2)$ | $r_{ijc,t+1} = \alpha_{ijc} + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_c / DD_{c,i}$ |
| II(c) | Rating $(\sigma^2)$ | $r_{ijc,t+1} = \alpha_c + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_i / DD_{\mathcal{R}}$ |
| II(d) | Fixed Effects $(\sigma_i^2)$ | $r_{ijc,t+1} = \alpha_{ijc} + \beta_c \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_i / DD_{\mathcal{R}}$ |
| III(a) | CAPM | $r_{ijc,t+1} = \alpha_{ijc} + \beta_{ijc} \bar{r}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_i / DD_{\mathcal{R}}$ |
| III(b) | Sector/CAPM | $r_{ijc,t+1} = \alpha_{ijc} + \beta_{1,ijc} \bar{r}_{c,t+1}$ $+ \beta_{2,ijc} \bar{r}_{cj,t+1} + u_{ijc,t+1}$ | $\lambda_i / DD_{\mathcal{R}}$ |
| IV | PCA | $r_{ijc,t+1} = \alpha_{ijc} + \boldsymbol{\beta}'_{ijc} \mathbf{f}_{c,t+1} + u_{ijc,t+1}$ | $\lambda_i / DD_{\mathcal{R}}$ |

Note: For the "Pooled" models the $c$ subscript on $\bar{r}_{c,t+1}, \bar{r}_{cj,t+1}, f_{c,t+1}$, and $\lambda_c$ is dropped.

For simulation of loss distributions, in addition to the return equations, we also need to specify the determination of the default thresholds. The first three models, I, II(a) and II(b), do not make use of credit rating information at the firm level. The Vasicek model treats all firms identically at the country level (for the two-country pooled model, *all* firms are treated identically), and so imposing within type homogeneity of distance to default is identical to within type homogeneity of default thresholds. Models II(a) and II(b) use within a given type homogeneity of default thresholds where firms are typed at the country level, or not at all for the pooled models. The default thresholds for the remaining models, labeled $\lambda_i / DD_{\mathcal{R}}$ in the Table 3, use the identifying restriction (9), namely that the distance to default is the same across all firms of a given rating. An estimate of $\lambda_i$ is obtained using (11) where estimates $\mu_i$ and $\sigma_{\xi_i}^2$ are obtained using (72) and (74), respectively.

We shall also consider simulation results using the alternative identification scheme for the default thresholds based on (8), which imposes the same threshold for all firms in a given rating. Further, to allow for direct comparisons, all models are calibrated to have the same expected loss within a cohort (or sample period).

## 9.3 Return Regression Results

The return regression parameters, estimated recursively using a 10-year rolling window, are summarized in Table 4. Note that these are all in-sample estimates. 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.

We focus for now on Panel A and shall return to Panels B through F of Table 4 below. The average pair-wise correlation of quarterly returns for the first ten years, 1988-1997, was 0.1933 for the U.S. firms in the sample, a much higher 0.6011 for the Japanese firms, and when all firms are pooled their average pair-wise correlation is very close to the U.S.-only sample at 0.1937, suggesting that the average pair-wise correlation between U.S. and Japanese firms is relatively low.[33]

The different model specifications are able to control for different degrees of cross firm return correlations. Considering first the U.S. and Japan pooled results, the average pair-wise correlation of residuals for the whole portfolio is around 0.022 for the Vasicek, fixed effect and single factor CAPM model. Adding an industry factor reduces that residual correlation to 0.0147, and the PCA model leaves almost no residual correlation. In-sample goodness of fit across models as measured by $\bar{R}^2$ (not reported in the table) range from 0.135 for the Vasicek to 0.229 for the sector CAPM to 0.339 for the PCA model.

Staying with the pooled model, notice the high degree of residual correlation that remains for the Japanese firms, ranging from around 45% (fixed effect) to 39% (CAPM). The reason is simple: the "global" market weighted return, $\bar{r}_{t+1}$, is dominated by U.S. firms. Still, if we examine U.S. firms only, their average pair-wise correlation of residuals ranges from around 7% (sector CAPM model) to 9.5% (Vasicek and fixed effects models). The overall portfolio average is so low only because residuals from U.S. and Japanese firm regressions are uncorrelated or even negatively correlated.

Estimating the models separately for each country helps, and this is seen clearly in the last three columns of Table 4, Panel A. While the overall average pair-wise correlation of residuals is quite similar at around 1.5% to 2%, for Japanese firms it is reduced dramatically, from a range of 39% to 45% under the pooled specification to a range of 2% to 6% when estimated separately. Similarly for U.S. firms, the average pair-wise correlation is reduced from a range of 7% to 9.5% in the pooled approach to a range of 2% to 3.5% when estimated separately. Clearly geographic heterogeneity plays an important role.

The results reported in Table 4 also show the high degree of variability in the coefficient estimates that exists across firms. This is illustrated by Figure 2 where the empirical densities (smoothed histograms) of the firm betas based on the one factor or CAPM model (Model III(a)) are displayed separately for the two countries. Clearly, there are substantial differences across firms and across countries. Also, the estimates of the Japanese betas are much more tightly distributed around their mean than are U.S. betas. We see a similar pattern with the firm alphas. These results are line with our assumption in Section 8 that the parameters of the return equations across the two

---

[33]For 1988-1997, the average pairwise correlation of USD-denominated returns for Japanese firms of 0.6011 is slightly higher than the average correlation of Yen-denominated returns of 0.5520 due to the common currency adjustment. However, local currency returns for Japanese firms in our sample are still noticeably more correlated than those for U.S. firms. This pattern holds for the later periods as well.

countries can be viewed as draws from two different distributions.

Panels B through F in Table 4 show the recursive estimation 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 (though they show a steady decline for Japanese firms), but starting with the cohort of 1991-2000 (Panel D), average correlation for the portfolio drops to 0.1391. By this point Japanese returns have an average pair-wise correlation of under 50% and approach 40% for the last cohort, 1993-2002. 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.

Turning now to the different models, the basic pattern across models remains unchanged as we move down the table (and thus forward in time). Figures 3a (U.S.) and 3b (Japan) show the empirical density plots of firm betas for three time periods: years ending 1997, 2000 and 2002. We notice for both countries that the distribution of estimated betas has been shifting subtly to the left. While the dispersion of U.S. betas has not changed much over the course of these rolling windows, the Japanese distribution appears to be widening. The firm heterogeneity we seek to explore here is thus exhibiting some time variation as well.

Finally, throughout the analysis we have been assuming time invariant volatilities. While it is well know 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 of the ARCH effects on all firm return regressions in the case of Model III(a). In Table 5 we report the percentage of firm regressions in which the ARCH effects are significant at the 5% level. Focusing first on the pooled country results, in the case of U.S. firms the proportion is close to the nominal value for the first three periods. Thereafter, there is increasing evidence of the ARCH effects, though never more than 10% of the firm regressions. Japanese firm returns, on the other hand, exhibit stronger ARCH effects as we can reject the hypothesis of no ARCH effects for around 10% of firm regressions over the six cohorts. The results for the separate country models are quite similar. Overall, however, the evidence is not sufficiently overwhelming to motivate ARCH modeling across all firms, and we do not expect serious biases resulting from our assumption of time invariant volatilities.

## 9.4 Impact of Heterogeneity on Credit Losses

With the firm regressions and default thresholds in hand, we are now able to simulate out-of-sample and compute the resulting loss distributions for the different model specifications. Here we provide a brief account of how the simulations were carried out. A more detailed description can be found in Appendix B. Firm returns, as specified in (71), are simulated out-of-sample assuming that the systematic and idiosyncratic components are serially uncorrelated and independently distributed,

thus imposing the conditional independence. For simplicity we set LGD = 100%. Allowing for different degrees of parameter heterogeneity yields different firm returns. For each simulation experiment we run 200,000 replications.

### 9.4.1 Comparison to Asymptotic Results for Vasicek Model

We begin by comparing the simulated loss distributions for our finite-sized portfolio to the asymptotic portfolio results which are available for the Vasicek model. Of interest are loss volatility or unexpected loss (UL), given by (32), and various quantile or VaRs. The analytical expressions for VaRs for this homogeneous model can be found in Vasicek (2002).

Table 6 reports the loss simulation results for the "pooled" version of Model I for each of the six rolling windows A through F. These results are out-of-sample in that the models, fitted over a ten-year sample, are used to forecast firm losses for the year immediately after the estimation period. For example, the first row in Table 6 describes the losses forecast in 1998 using the model estimated over the sample period 1988-1997. For each year we report the portfolio expected default rate, $\hat{\pi}$, which is equal to expected loss under our assumption of no loss recovery and the average portfolio return correlation, $\rho$, given by the empirical analog of (27):

$$\hat{\rho} = \frac{\hat{\beta}^2 \hat{V}(\bar{r}_{t+1})}{\hat{\beta}^2 V(\bar{r}_{t+1}) + \hat{\sigma}_u^2},$$
(78)

where $\hat{\beta}$ and $\hat{\sigma}_u^2$ are computed recursively using Model I. The estimated variance of $\bar{r}_{t+1}$, $\hat{V}(\bar{r}_{t+1})$, is computed from the aggregate returns, $\bar{r}_{t+1}$, using a rolling 10-year observation window. For the size of the portfolio we report the total number of firms, $N_t$, and the effective number of equal-sized exposures, $N_t^* = \left(\sum_{i=1}^{N_t} w_{it}^2\right)^{-1}$, where $w_{it}$ are the exposure weights. Table 6 also reports the asymptotic and simulated UL and VaRs, as well as their differences denoted as "granularity."

Looking across the first row of Table 6 we see that our two-country portfolio of 839 firms, with an effective number of 638 equal-sized exposures, is relatively close to an asymptotically diversified portfolio. Simulated UL is 1.47%, only 7bp above the asymptotic result. Similarly for the three quantiles we obtain 99.0%, 99.5% and 99.9% VaR, the last corresponds to the loss calibration level of the New Basel Accord (BCBS, 2004), the simulated loss levels are never far from though always above their asymptotic counterpart, as is to be expected. For instance, simulated 99.9% VaR is 12.05% of total portfolio notional, just 23bp above the level achievable with an infinitely large portfolio. Looking down the table, it is clear that the two-country portfolios for later cohorts are also close to an asymptotically diversified portfolio.

Introduction of credit rating information in the Vasicek model affects the losses through changes to the default correlation, $\rho^*$, even though the pair-wise return correlations continue to be the same across all firms. From (31) and Figure 1 we know that for any given value of return correlation, $\rho$, default correlation, $\rho^*$, is increasing in the unconditional probability of default, $\pi$, so long as

36

$\pi < 0.5$. This can be seen clearly by computing the default correlation by credit rating, $\rho_{\mathcal{R}}^*$, given an estimate of a rating specific default rate, $\hat{\pi}_{\mathcal{R}}$, and the same distance to default by rating identifying restriction, $DD_{\mathcal{R}}$, given by (9). These correlations can be easily computed using (18), repeated here for convenience:

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

where $\pi_{i,t+1} = E\left(z_{i,t+1}\right)$ is firm $i^{th}$ default probability over the period $t$ to $t+1$. Under the double Gaussian assumption, $E\left(z_{i,t+1}z_{j,t+1}\right)$ is simply

$$E\left(z_{i,t+1}z_{j,t+1}\right) = \Phi_2\left[\Phi^{-1}\left(\pi_{i,t+1}\right), \Phi^{-1}\left(\pi_{j,t+1}\right), \rho_{ij}\right], \tag{79}$$

where $\Phi_2[\cdot]$ is the bivariate standard normal CDF, and $\rho_{ij}$ is the return correlation between firm $i$ and $j$ assumed fixed at $\rho$ under the Vasicek's model. In Panel A of Table 7 we report default correlations by credit rating for 2003 based on estimates using the last rolling sample window, 1993-2002.[34] The top two ratings, $\mathcal{AAA}$ & $\mathcal{AA}$, are combined because we impose a minimum unconditional default probability of 0.001% per annum; see also the discussion in Section 9.1.

The strong differentiation of the default correlation by credit rating, $\hat{\rho}_{\mathcal{R}}^*$, suggests that this rating is indeed an important characterization of default heterogeneity, even while imposing strict homogeneity of return correlations. Put differently, a high default rate for a low credit rating, say $\mathcal{B}$, is associated with high default correlation, and *vice versa*. Gordy (2000) reports similar steeply increasing default probabilities as one descends the credit rating scale.

Default correlations across credit ratings can also be similarly computed, and these are summarized in Panel B of Table 7. The diagonal entries are just the within-rating default correlation presented in Panel A. As such, the matrix in Panel B should be thought of as a matrix of correlations rather than a correlation matrix. We notice that the pair-wise default correlations between high-yields firms, i.e. with ratings lower than $\mathcal{BBB}$, can be substantial.

### 9.4.2 Loss Simulation Results

Although most of our loss simulations to be reported below are calibrated to have the same EL across all models, we start by examining the impact of a simple source of heterogeneity, namely firm fixed effects, on expected and unexpected losses. Recall from the theoretical results in Section 6.3 that the introduction of heterogeneity in $a_i = (\lambda_i - \mu_i)/\sigma$, namely for $\sigma_v^2 > 0$,[35] resulted in increased expected and decreased unexpected losses (once EL is controlled for). Empirically one may introduce this heterogeneity through a firm fixed effect which impacts $\mu_i$ and therefore $a_i$. The results are displayed in Table 8, where Panel A shows the impact on EL while Panel B controls for EL to be the same across the two model specifications and shows the impact on UL. For the

---

[34]We do not report the results for previous years as they are qualitatively very similar.

[35]Recall $a_i = a + v_i$, $\quad v_i \sim iid\ N(0, \sigma_v^2)$.

first cohort, 1998, the simulated EL increases from 1.23% to 1.72%, as predicted by theory. Panel B shows that UL declines from 1.47% to 1.39%, again as predicted. This pattern holds true for all six years or cohorts.

Table 9 reports the loss simulation results for each of the six rolling windows. We proceed by discussing in detail the loss simulations for the first period and then draw comparisons across years. In addition to the model number and name in Table 9, we provide the EL calibration using the default threshold $\lambda$ and distance to default, $DD$. For example, the basic Vasicek model has the same threshold $\lambda$ and the same distance to default, $DD$, for all firms. For Model II(a) we add fixed effects with a common firm error variance, which requires $DD$ to be firm specific while keeping $\lambda$ fixed across firms.

For each year we report the first four simulated moments of the loss distribution (note that the first moment or expected loss is the same across all models by construction) as well as three commonly reported quantiles: 99.0%, 99.5% and 99.9%. We also calculated expected shortfall; the results are qualitatively no different, and so we report here only the VaR results. The first set of columns is for the pooled specification while the second set is for the country specific models, analogous to the in-sample regression results in Table 4.

The fully homogeneous model of Vasicek (Model I), generates the most extreme losses and has the largest unexpected losses. Allowing for firm fixed effects but keeping firm error variances the same, Model II(a), results in only a small reduction in risk (whilst controlling for expected losses). UL drops slightly from 1.47% to 1.39%, the resulting loss distribution is somewhat less skewed (2.8 vs. 3.1) and fat-tailed (kurtosis of 16.9 vs. 19.5), and risk as measured by VaR is a bit lower as well. For instance 99.9% VaR drops from 12.05% of portfolio notional to 11.14%.

Note that a bank complying with an 8% minimum capital requirement would be insolvent at the 99.9% VaR under these two model specifications. To be sure, the assumption of 100% LGD made here is unrealistic; average empirical LGDs are around half that. A quick glance at the following rows suggests that allowing for parameter heterogeneity beyond fixed effects, for instance by just relaxing the error variances to be firm specific, i.e. Model II(b), results in 99.9% VaR losses of around 5%. Indeed this source of heterogeneity turns out to be quite important. UL is reduced by about 70%, dropping to 0.80% from 1.39%. The shape of the overall distribution is less extreme as both skewness and kurtosis decline substantially. Finally, allowing for firm specific error variances results in 99.9% VaR of 5.11%, less than half of the value obtained for Model II(a), 11.14%.

The above simulations do not make use of firm rating information.[36] Models II(c) and II(d) examine the impact or value of firm credit ratings. Model II(c) simply adds ratings in the form of a rating-specific default threshold $\lambda$ and distance to default $DD$ to the Vasicek specification in Model I. For the pooled model we go from one (two for the country-specific models) to seven (fourteen)

---

[36] To be sure, the overall portfolio EL is calibrated using default rates by credit rating, but so far we have not used this information at a firm level.

$\lambda's$, along the lines outlined in Section 3. It is clear that ratings capture an important source of firm heterogeneity, at least as it pertains to credit risk. Comparing Model II(c) to Model I, UL drops by about one-third from 1.47% to 1.07% while 99.9% VaR is reduced by nearly 80% from 12.05% to 6.72%. 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. It is no accident that credit ratings, whether internal or external to the bank, are one of the important cornerstones of the New Basel Accord.

Compared to Model II(c), Model II(d) allows for fixed effects and firm specific error variances. Model II(d) can also be compared to Model II(b) with rating information added. Although UL is a bit higher for Model II(c) as compared to Model II(b), 0.84% vs. 0.80%, VaR numbers are lower for all three VaR levels. For instance, at the 99.9% level, VaR is 4.81% compared to 5.11% for Model II(b). Indeed, a quick glance across all models indicates that VaR is lowest for Model II(d).

Models III(a) and (b) add heterogeneous slopes to Model II(d), with Model III(b) also adding an industry return factor. Overall the difference in risk is small. UL and VaR increase somewhat with the addition of these parameters. Adding an industry factor above and beyond a market factor with heterogeneous slopes makes little difference to the resulting loss distribution as is clear by comparing Models III(a) and III(b). UL is nearly the same, 0.86% vs. 0.88%, as are VaR levels: 5.56% vs. 5.58% at the 99.9% VaR level.

Finally, the principal components model IV generates UL results which are similar to Model II(c), which is Model I with ratings information, namely 1.08% vs. 1.07%. VaR, however, is higher. For instance, 99.9% VaR is 7.69%, quite close to the 8% minimum capital requirement, compared to 6.72% for Model II(c). In this way Model IV also generates tail losses which are higher than Models III(a) and III(b). Although the PCA model performs best on an in-sample basis, the recursive out-of-sample simulations do not seem to support the in-sample evidence.

Allowing for geographic heterogeneity in the simple Vasicek case helps as can be seen by the last set of columns labeled "U.S. & Japan Modeled Separately." For the Vasicek model this amounts to doubling the number of parameters as there is one set for each country. UL drops by about 14% from 1.47% to 1.29%, skewness and kurtosis both decline, as does VaR. For instance, 99.9% VaR declines by nearly 20% from 12.05% tp 10.14%.

In general, however, pooled and separate country models generate similar loss distributions. Modeling the U.S. and Japan separately usually results in lower risk for Models I and II(a). Once firm specific slopes are allowed, as in Models III(a) and (b), VaR is actually slightly higher for the separate country model, although the second through fourth moments are the same. For instance, in the case of the basic CAPM model, Model III(a), 99.9% VaR increases from 5.56% to 5.75%. A similar pattern can be seen when adding an industry factor in Model III(b). Model IV, however, shows risk reduction by allowing for country heterogeneity. Moreover, the results for the last period shown in Panel F, Table 9, indicates that even for the simple Vasicek model, Model I, pooling need

not always increase risk: 99.9% VaR is 17.47% for the pooled model but 17.81% for the country specific model. Indeed for this last period, allowing the parameters to vary by country increases VaR for all models save the last one, Model IV. Some of these seemingly adverse results could be due to parameter instability and the associated estimation errors. Nevertheless, the model ranking is robust to comparisons over time, i.e. comparing Panels A through F. **MHP:**Overall it seems that allowing for country specific factor loadings is more important than requiring different country models to have their own specific factors.

The difference in the loss distributions across models can also be seen in Figure 4 where we show the loss densities for 2003 across the different specifications (separate country models). There are clearly important differences across the models. Relative to the Vasicek model which has only three parameters per country $(\alpha_c, \beta_c, \sigma_c)$, the loss densities of the other models are not all that different. Adding just fixed effects, Model II(a), does not seem to change the shape of the loss distribution. However, once the firm error variances are allowed to differ, Model II(b), the distributional shape changes dramatically. Interestingly, Model II(c), which simply adds credit rating information to the Model I specification (i.e. it adds just seven parameters per country) yields a loss distribution which is remarkably similar to ones generated by the fully heterogeneous model specification. Credit ratings seem indeed to be a useful and informative summary statistic for firm-level credit risk.

Moving down the panels in Table 9 we notice that the portfolio is getting riskier over time; expected loss rises every year. If we compare value-at-risk, say at the 99.5%, for a model, say the fixed effects model with firm specific error variances and ratings (Model II(d)) applied to the two countries separately, we see that VaR increases from 3.84% in 1998 to 5.56% in 2000 to 6.91% in 2003.

When considering credit risk, there are two types of parameter heterogeneity that should be taken into account: firm returns and default thresholds (or distance to default). What is the impact of one vs. the other? Table 10 seeks to address this question, where in Panel A we have repeated several model results from Panel F in Table 9, namely the loss simulation results for the final year under the basic same distance to default identification assumption given by (9). Panel B imposes the other identifying assumption, namely the same default threshold $\lambda$ as in (8). First, note that by construction the two identification procedures are identical for Models I and II(c). Focusing on the pooled models, we find that imposing (9) in Panel A generates levels of 99.9% value-at-risk that are 4% (for Model II(a)) to 13% (for Model IV) higher than those generated by (8) in Panel B. For instance, under (8) Model IV yields 99.9% VaR of 9.83% as seen in Panel B, far less than 99.9% VaR of 11.15% in Panel A. We are currently exploring more fully these results, but initial closer examination reveals that imposing same $DD$, (9), generates more extreme tail events for investment grade firms than does imposing same $\lambda$, (8), while the differences for speculative grade firms are more modest.

What is clear from the discussion is that there is a rich and complex interaction between the

underlying model parameters and the resulting loss distributions. Some of this complexity can be summarized more simply using the default correlation $\rho_{ij}^*$ between any two firms $i$ and $j$, as computed in Section 9.4.1 where we discussed the default correlation by credit rating, $\rho_{\mathcal{R}}^*$. We can conduct similar analysis across a broader set of models that use the same $DD$ identifying assumption, and these results are summarized in Table 11 for the year, 2003. We report the portfolio EL for the year, 3.26%, which is the same across all models, the average return correlation computed using (16), the average default correlation using (79) as well as portfolio loss volatility, UL, taken from Table 9. Note that for the first two models, I and II(c), parameter restrictions result in only one return correlation $\rho$ for all firms. Through the introduction of ratings information and consequently heterogeneity in unconditional default probabilities, we notice a sharp reduction in average default correlation, $\bar{\rho}^*$, from Model I, 1.81% to II(c), 0.26%. This trend continues when we allow in addition for firm fixed effects and firm-specific error variances, Model II(d), where $\overline{\hat{\rho}}^* = 0.18\%$. This reduction in average default correlation is accompanied by a reduction in UL, as seen in the last column of Table 11.

Default correlation and UL exhibit a U-shaped relationship across models. Both risk measures are highest for the simplest model, Model I, lowest for a model of intermediate complexity and heterogeneity, Model II(d), and then rise again as we increase the heterogeneity in the conditional mean specification. For instance, Model III(b), the two-factor model, has a higher average return correlation than Model II(d), 10.35% vs. 9.06%, higher average default correlation, 0.33% vs. 0.18%, and also higher portfolio UL, 1.28% vs. 1.16%. Model IV, the PCA specification, is higher still for all three of these measures. By allowing for firm specific error variances and making use of credit ratings, Model II(d) allows for a high degree of heterogeneity in the default condition and thereby allows for significant diversification of the idiosyncratic risk. The loss simulations are, of course, conducted under the assumption of conditional independence, but Model II(d) imposes a great deal of homogeneity on the conditional mean specification where the return correlation is captured. As a means of comparison, the in-sample average pair-wise return correlations, on which the parameter estimates for all models are based, is 15.45% (see Table 4, Panel F). The two-factor model, III(b), comes closer to this value at 10.35% than does the simple firm fixed effect model, II(d), at 9.06%. So the low average default correlation and portfolio UL generated by Model II(d) reflects, at least to some degree, the failure to properly account for all of the systematic risk, as summarized by the average return correlation. In this way Model II(d) can be said to be overly optimistic from the point of view of portfolio credit risk.

# 10   Concluding Remarks

In this paper we have considered a simple model of credit risk and derived the limit distribution of losses under different assumptions regarding the structure of systematic risk and the scope of

exposure or firm heterogeneity. The analytical and simulation results point to some interesting conclusions. Theory indicates that if the firm parameters are heterogeneous but come from a common distribution, asymptotically there is no scope for further risk reduction through active portfolio management. However, if the firm parameters come from different distributions, say for different sectors or countries, then further risk reduction is possible, even asymptotically, by changing the portfolio weights. 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 particularly important? Firm fixed effects do not seem very important, particularly as compared to allowing for variation in return error variances across firms. This can be accomplished either by allowing each firm to have its own error variance or by appropriate grouping such as by credit rating. Indeed going from one type to effectively seven, one for each credit rating, with seven unconditional default probabilities, seems to capture much of the relevant heterogeneity. Unfortunately there is insufficient data to meaningfully differentiate the unconditional default probabilities by rating *and* by country.

Allowing for flexible factor sensitivities appears to be of secondary importance. Risk does not seem to change much beyond Model III(a), the basic one-factor model. However, this is misleading. Recall that all of the loss simulations are done under the assumption of conditional independence. If this assumption 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, and here the addition of industry return factors and the concomitant heterogeneous factor sensitivities, e.g. Model III(b), can become quite important.

The average pair-wise cross-sectional correlation of residuals are similar across the first set of models which all have just a single factor. However, starting with Model III(b), those average correlations decline (from about 23% to about 15%), and they become negligible, by construction, with the PCA model. Indeed the risk forecast by those last three models (Model III(a) and (b), Model IV) increases as the average pairwise correlation of residuals decreases. Put differently, while Model IV is conditionally independent on an in-sample basis, Models III(a) and III(b) are not. So long as on an out-of-sample basis Model IV is still closer to conditional independence than Models III(a) and III(b), the latter will generate lower, risk forecasts than the former. Measuring and evaluating out-of-sample conditional dependence is an important topic which requires further investigation.

Finally, the differences in pooled versus country specific results suggest that further sub-dividing the firm return specification and error variances by country matters less. Indeed, the differences within a given model in the pooled approach compared with modeling the two countries separately

are generally smaller than the differences across models for a given approach. Thus allowing for country specific factor loadings is important, but requiring country specific factors is less so.

# A  Principal Components Model

For the principal components model where

$$r_{i,t+1} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{f}_{t+1} + u_{i,t+1}, \ t = 1, ..., T$$

the estimates of $\mathbf{f}_{t+1}$ can be computed by application of principal component techniques to the standardized returns defined by $(r_{i,t+1} - \bar{r}_i)/\hat{\sigma}_i$, where $\bar{r}_i = \sum_{t=1}^{T} r_{i,t+1}/T$, and $\hat{\sigma}_i = \sum_{t=1}^{T} (r_{i,t+1} - \bar{r}_i)^2/(T-1)$. The number of factors, $m$, can be selected using the Bai and Ng (2002) procedure.

Denote the number of selected factors by $\hat{m}$ and estimated factors by $\hat{\mathbf{f}}_{t+1}$, and suppose that they are subject to the orthonormalization restrictions such that[37]

$$(T)^{-1} \sum_{t=1}^{T} \hat{\mathbf{f}}_{t+1} \hat{\mathbf{f}}_{t+1}' = \mathbf{I}_{\hat{m}}, \ T^{-1} \sum_{t=1}^{T} \hat{\mathbf{f}}_{t+1} = \mathbf{0}. \tag{80}$$

The estimates of $\alpha_i$ and $\boldsymbol{\beta}_i$ will be given by

$$\hat{\alpha}_i = T^{-1} \sum_{t=1}^{T} r_{i,t+1} = \bar{r}_i, \qquad \hat{\boldsymbol{\beta}}_i = T^{-1} \sum_{t=1}^{T} (r_{i,t+1}) \hat{\mathbf{f}}_{t+1},$$

and

$$\hat{\sigma}_i^2 = (T - \hat{m} - 1)^{-1} \sum_{t=1}^{T} \left( r_{i,t+1} - \hat{a}_i - \hat{\boldsymbol{\beta}}_i' \hat{\mathbf{f}}_{t+1} \right)^2. \tag{81}$$

# B  Simulation of Returns and Associated Loss Distributions

To simulate $h-$period ahead individual firm returns, $R_{i,t+h} \equiv \sum_{l=1}^{h} r_{i,t+l}$, according to (71) we need to simulate $\mathbf{f}_{t+l}$, and $u_{i,t+l}$, for $l = 1, 2, .., h$. We assume that the common factors are serially uncorrelated. This seems justified in our application where market returns are only very weakly autocorrelated. Taking first the case of a single market factor, suppose that

$$\bar{r}_{t+l} \sim iidN(\bar{r}, s_{\bar{r}}^2),$$

$$\bar{r} = T^{-1} \sum_{t=1}^{T} \bar{r}_t, \qquad s_{\bar{r}}^2 = (T-1)^{-1} \sum_{t=1}^{T} (\bar{r}_t - \bar{r})^2,$$

and

$$u_{i,t+l} \sim iidN(0, \hat{\sigma}_i^2).$$

Then return for the $i^{th}$ firm can be simulated as

$$r_{i,t+l}^{(s)} = \hat{\alpha}_i + \hat{\beta}_i \bar{r}_{t+l}^{(s)} + \hat{\sigma}_i \varepsilon_{i,t+l}^{(s)},$$

---

[37] For computational convenience, we actually impose the restriction that the *sample* covariance matrix $(T-1)^{-1} \sum_{t=1}^{T} \hat{\mathbf{f}}_{t+1} \hat{\mathbf{f}}_{t+1}' = \mathbf{I}_{\hat{m}}$.

$$\bar{r}_{t+l}^{(s)} \sim iidN(\bar{r}, s_{\bar{r}}^2), \ \varepsilon_{i,t+l}^{(s)} \sim iidN(0, 1)$$

where $s = 1, 2, .., S$ are the replications.

For the principal components model where the factors satisfy the orthonormalization restrictions (80), returns for the $i^{th}$ firm can be simulated as

$$r_{i,t+l}^{(s)} = \hat{\alpha}_i + \boldsymbol{\beta}_i' \mathbf{f}_{t+l}^{(s)} + \hat{\sigma}_i \varepsilon_{i,t+l}^{(s)},$$

$$\mathbf{f}_{t+l}^{(s)} \sim iidN(\mathbf{0}, \mathbf{I}_{\hat{m}}), \ \varepsilon_{i,t+l}^{(s)} \sim iidN(0, 1).$$

For the more general multi-factor case, let $\boldsymbol{\mu}_f$ denote the mean and $\boldsymbol{\Sigma}_f$ the covariance matrix of $\mathbf{f}_{t+l}$. Then returns for the $i^{th}$ firm can be simulated as

$$r_{i,t+l}^{(s)} = \hat{\alpha}_i + \boldsymbol{\beta}_i' \mathbf{f}_{t+l}^{(s)} + \hat{\sigma}_i \varepsilon_{i,t+l}^{(s)},$$

$$\mathbf{f}_{t+l}^{(s)} \sim iidN(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f), \ \varepsilon_{i,t+l}^{(s)} \sim iidN(0, 1).$$

Portfolio loss at horizon $h$ can now be simulated using

$$l_{t+h}^{(s)} = V_p \left[ \sum_{i=1}^{N} w_i I \left( \hat{\lambda}_{i,t+h} - R_{i,t+h}^{(s)} \right) \right],$$

where $R_{i,t+h}^{(s)} \equiv \sum_{l=1}^{h} r_{i,t+l}^{(s)}$ is the $h$-period cumulative return, $\hat{\lambda}_{i,t+h}$ is the $h$-period default return threshold, $V_p$ is the face value of the whole portfolio (e.g. \$1bn) and $w_i$ is the fraction of exposure to obligor $i$. We assume for simplicity that defaulted instruments have no recovery value. Simulated expected loss due to default is given by

$$\bar{l}_{S,t+h} = \frac{1}{S} \sum_{s=1}^{S} l_{t+h}^{(s)}.$$

The higher order moments of the loss distribution can be similarly simulated.

# References

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

[2] Altman, Edward I., Brooks Brady, Andrea Resti and Andrea Sironi (2003). "The Link between Default and Recovery Rates: Implications for Credit Risk Models and Procyclicality." forthcoming, *Journal of Business*.

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

[4] Basel Committee on Banking Supervision (2004). "International Convergence of Capital Measurement and Capital Standards: A Revised Framework." <http://www.bis.org/publ/bcbs107.htm>, June.

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

[6] Crouhy, Michel, Dan Galai, and Robert Mark (2000). "A Comparative Analysis of Current Credit Risk Models." *Journal of Banking & Finance* 24, 59-117.

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

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

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

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

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

[12] Duffie, Darrell and Ke Wang (2004). "Multi-Period Corporate Failure Prediction with Stochastic Covariates." Working Paper, Stanford Graduate School of Business.

[13] 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.

[14] Engle, Robert F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50, 987-1008.

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

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

[17] Frye, Jon (2000), "Depressing Recoveries," Risk 13, April, 91-94.

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

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

[20] Gordy, Michael B. (2000). "A Comparative Anatomy of Credit Risk Models." *Journal of Banking & Finance* 24 (1-2), 119-149.

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

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

[23] 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.

[24] Hu, Y-T and William Perraudin (2002). "The Dependence of Recovery Rates and Defaults." CEPR working paper.

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

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

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

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

[29] Lawrence, Edward C., and Nasser Arshadi. (1995), "A Multinomial Logit Analysis of Problem Loan Resolution Choices in Banking." *Journal of Money, Credit and Banking* 27, 202-216.

[30] 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.

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

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

[33] Mella-Barral, Pierre, and William Perraudin. (1997). "Strategic Debt Service." *Journal of Finance* 52, 531-556.

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

[35] Pesaran, M.Hashem, Til Schuermann, Björn-Jakob Treutler and Scott M. Weiner (2004) "Macroeconomic Dynamics and Credit Risk: A Global Perspective." forthcoming, *Journal of Money, Credit and Banking*; available as Wharton Financial Institutions Center Working Paper #03-13B.

[36] 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.

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

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

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

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

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

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

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

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

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

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

| | Sample Window | Average Pair-wise Correlation of Returns | | | Model Specifications | | Average Pair-wise Correlation of Residuals | | | | | |
| | | | | | | | US & Japan Pooled | | | US & Japan Modeled Separately | | |
| | | US&JP | US | JP | | | US&JP | US | JP | US&JP | US | JP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A** | **1988-1997** | 0.1937 | 0.1933 | 0.6011 | I | Vasicek | 0.0222 | 0.0951 | 0.4217 | 0.0239 | 0.0365 | 0.0607 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0222 | 0.0951 | 0.4217 | 0.0239 | 0.0365 | 0.0607 |
| | **Sample Size** | 839 | 628 | 211 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0228 | 0.0789 | 0.4490 | 0.0232 | 0.0361 | 0.0606 |
| | | | | | III(a) | CAPM | 0.0218 | 0.0797 | 0.3868 | 0.0242 | 0.0374 | 0.0639 |
| | | | | | III(b) | Sector CAPM | 0.0147 | 0.0711 | 0.3869 | 0.0142 | 0.0221 | 0.0227 |
| | | | | | IV | PCA | -0.0001 | 0.0016 | 0.0037 | 0.0000 | 0.0005 | -0.0045 |
| **Panel B** | **1989-1998** | 0.2150 | 0.2114 | 0.5913 | I | Vasicek | 0.0235 | 0.0940 | 0.3793 | 0.0249 | 0.0440 | 0.0633 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0235 | 0.0940 | 0.3793 | 0.0249 | 0.0440 | 0.0633 |
| | **Sample Size** | 854 | 633 | 221 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0241 | 0.0802 | 0.4059 | 0.0246 | 0.0438 | 0.0629 |
| | | | | | III(a) | CAPM | 0.0237 | 0.0825 | 0.3553 | 0.0255 | 0.0456 | 0.0666 |
| | | | | | III(b) | Sector CAPM | 0.0177 | 0.0773 | 0.3582 | 0.0193 | 0.0324 | 0.0258 |
| | | | | | IV | PCA | 0.0003 | 0.0011 | 0.0028 | 0.0000 | 0.0004 | -0.0041 |
| **Panel C** | **1990-1999** | 0.2097 | 0.2237 | 0.5666 | I | Vasicek | 0.0533 | 0.1282 | 0.3917 | 0.0627 | 0.0731 | 0.1391 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0533 | 0.1282 | 0.3917 | 0.0627 | 0.0731 | 0.1391 |
| | **Sample Size** | 842 | 613 | 229 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0543 | 0.1187 | 0.4112 | 0.0620 | 0.0736 | 0.1392 |
| | | | | | III(a) | CAPM | 0.0568 | 0.1246 | 0.3807 | 0.0661 | 0.0778 | 0.1473 |
| | | | | | III(b) | Sector CAPM | 0.0419 | 0.1168 | 0.3817 | 0.0417 | 0.0621 | 0.0390 |
| | | | | | IV | PCA | 0.0000 | 0.0005 | 0.0010 | -0.0003 | -0.0001 | -0.0036 |
| **Panel D** | **1991-2000** | 0.1391 | 0.1691 | 0.4638 | I | Vasicek | 0.0533 | 0.1207 | 0.3427 | 0.0633 | 0.0783 | 0.1359 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0533 | 0.1207 | 0.3427 | 0.0633 | 0.0783 | 0.1359 |
| | **Sample Size** | 816 | 588 | 228 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0540 | 0.1142 | 0.3589 | 0.0607 | 0.0788 | 0.1372 |
| | | | | | III(a) | CAPM | 0.0559 | 0.1206 | 0.3442 | 0.0649 | 0.0821 | 0.1428 |
| | | | | | III(b) | Sector CAPM | 0.0379 | 0.1039 | 0.3470 | 0.0392 | 0.0615 | 0.0427 |
| | | | | | IV | PCA | 0.0000 | 0.0018 | 0.0030 | 0.0002 | 0.0008 | -0.0041 |
| **Panel E** | **1992-2001** | 0.1309 | 0.1633 | 0.4411 | I | Vasicek | 0.0482 | 0.1132 | 0.3282 | 0.0578 | 0.0740 | 0.1277 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0482 | 0.1132 | 0.3282 | 0.0578 | 0.0740 | 0.1277 |
| | **Sample Size** | 811 | 585 | 226 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0490 | 0.1074 | 0.3436 | 0.0545 | 0.0745 | 0.1290 |
| | | | | | III(a) | CAPM | 0.0501 | 0.1135 | 0.3324 | 0.0583 | 0.0772 | 0.1339 |
| | | | | | III(b) | Sector CAPM | 0.0378 | 0.1077 | 0.3361 | 0.0381 | 0.0624 | 0.0421 |
| | | | | | IV | PCA | -0.0007 | 0.0001 | 0.0009 | -0.0004 | -0.0003 | -0.0042 |
| **Panel F** | **1993-2002** | 0.1545 | 0.1999 | 0.4191 | I | Vasicek | 0.0549 | 0.1098 | 0.3332 | 0.0628 | 0.0772 | 0.1254 |
| | | | | | II(a) | Fixed Effect ($\sigma^2$) | 0.0549 | 0.1098 | 0.3332 | 0.0628 | 0.0772 | 0.1254 |
| | **Sample Size** | 818 | 600 | 218 | II(b) | Fixed Effect ($\sigma^2_i$) | 0.0556 | 0.1097 | 0.3377 | 0.0600 | 0.0780 | 0.1270 |
| | | | | | III(a) | CAPM | 0.0569 | 0.1157 | 0.3488 | 0.0641 | 0.0811 | 0.1324 |
| | | | | | III(b) | Sector CAPM | 0.0439 | 0.1099 | 0.3543 | 0.0426 | 0.0658 | 0.0391 |
| | | | | | IV | PCA | -0.0008 | -0.0006 | 0.0001 | -0.0007 | -0.0009 | -0.0043 |

**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. The table presents estimation results for the "pooled" country models in the first set of columns and the "separate" country models in the second set of columns. Portfolio determination and sample construction are discussed in Section 9.1. Specification of the return models is discussed in Section 9.2 (see Table 3 for further detail). The data source for returns of U.S. firms is CRSP. The source for Japanese firms is Datastream. Yen-denominated Japanese returns are converted to USD-denominated returns by subtracting the percentage change in the Yen/USD exchange rate.

## Table 5
## Testing for ARCH Effects in Model III(a) CAPM

| | % of Firms with Significant ARCH Effects Nominal Significance Level = 5% | | | | | |
| Sample | US & Japan Pooled | | | US & Japan Modeled Separately | | |
| Window | US&JP | US | JP | US&JP | US | JP |
|---|---|---|---|---|---|---|
| **1988-1997** | 6.56% | 4.94% | 11.37% | 5.24% | 4.94% | 6.16% |
| **1989-1998** | 6.32% | 4.27% | 12.22% | 5.85% | 5.06% | 8.14% |
| **1990-1999** | 7.72% | 5.55% | 13.54% | 6.29% | 4.89% | 10.04% |
| **1991-2000** | 9.56% | 8.33% | 12.72% | 10.05% | 9.35% | 11.84% |
| **1992-2001** | 9.62% | 8.55% | 12.39% | 10.11% | 9.40% | 11.95% |
| **1993-2002** | 9.90% | 9.33% | 11.47% | 9.41% | 9.00% | 10.55% |

**Note:** The table reports the result of tests for GARCH effects conducted using recursive estimation results for Model III(a) - "CAPM" - presented in Table 4. Tests are carried out using Engle's (1982) Lagrange Multiplier Test. The figure reports the percentage of firms that have ARCH effects at the 5% significance level.

**Table 6**
**Analytical and Simulated Loss Distribution Results for Vasicek Model**
**Based on U.S. & Japan Pooled Return Specifications**

| Simulation | | | Portfolio Size | | UL | | | 99.0% VaR | | | 99.5% VaR | | | 99.9% VaR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | $\hat{\pi}$ | $\hat{\rho}$ | N | N* | Anal. | Simul. | Granul. | Anal. | Simul. | Granul. | Anal. | Simul. | Granul. | Anal. | Simul. | Granul. |
| 1998 | 1.23% | 13.83% | 839 | 638 | 1.40% | 1.47% | 0.07% | 6.81% | 7.00% | 0.19% | 8.23% | 8.48% | 0.26% | 11.82% | 12.05% | 0.23% |
| 1999 | 1.60% | 15.69% | 854 | 635 | 1.89% | 1.94% | 0.05% | 9.17% | 9.30% | 0.13% | 11.07% | 11.13% | 0.06% | 15.84% | 15.91% | 0.07% |
| 2000 | 2.10% | 13.71% | 842 | 613 | 2.14% | 2.22% | 0.08% | 10.35% | 10.65% | 0.30% | 12.25% | 12.59% | 0.33% | 16.92% | 17.34% | 0.42% |
| 2001 | 2.27% | 8.00% | 816 | 619 | 1.65% | 1.75% | 0.10% | 8.08% | 8.39% | 0.31% | 9.24% | 9.66% | 0.42% | 12.01% | 12.72% | 0.71% |
| 2002 | 2.73% | 7.71% | 811 | 642 | 1.88% | 1.98% | 0.10% | 9.22% | 9.48% | 0.26% | 10.47% | 10.74% | 0.27% | 13.42% | 13.71% | 0.29% |
| 2003 | 3.26% | 9.20% | 818 | 654 | 2.38% | 2.48% | 0.10% | 11.61% | 11.92% | 0.31% | 13.23% | 13.59% | 0.36% | 17.06% | 17.47% | 0.41% |

**Note:** This table reports comparisons of our simulated loss distributions for the U.S. & Japan Pooled Vasicek model (Model I) with analytical results for an infinitely large portfolio. $\hat{\pi}$ is the portfolio EL and $\hat{\rho}$ is estimated using equation (76). N is the number of obligors in our portfolio. $N^* = ( \Sigma w_i^2 )^{-1}$ is the equivalent number of equal weighted exposures. Analytical Unexpected Loss (UL) and Value-at-Risk (VaR) are computed using the expressions given in Vasicek (2002). Granularity is Simulated minus Analytical UL. The simulation procedure is discussed in Appendix B. All simulations are carried out using 200,000 replications.

# Table 7
## Simulated (Out-of-Sample) Default Correlations for 2003
## Based on Model II(c) for U.S. & Japan Pooled Return Specification

**Panel A:**      Default Probabilities (in basis points) and Correlations by Credit Rating

|  | $\hat{\pi}_R$ | $\hat{\rho}$ | $\hat{\rho}_R^*$ |
|---|---|---|---|
| **AAA & AA** | 0.10 | 9.20% | 0.004% |
| **A** | 0.58 | 9.20% | 0.018% |
| **BBB** | 10.59 | 9.20% | 0.170% |
| **BB** | 63.03 | 9.20% | 0.615% |
| **B** | 542.88 | 9.20% | 2.437% |
| **CCC** | 4,977.60 | 9.20% | 5.865% |

**Panel B:**      Default Correlations by Credit Rating

|  | **AAA & AA** | **A** | **BBB** | **BB** | **B** | **CCC** |
|---|---|---|---|---|---|---|
| **AAA & AA** | 0.004% | - | - | - | - | - |
| **A** | 0.009% | 0.018% | - | - | - | - |
| **BBB** | 0.026% | 0.054% | 0.170% | - | - | - |
| **BB** | 0.048% | 0.100% | 0.320% | 0.615% | - | - |
| **B** | 0.086% | 0.183% | 0.607% | 1.199% | 2.437% | - |
| **CCC** | 0.102% | 0.224% | 0.792% | 1.636% | 3.560% | 5.865% |

**Note:** This table reports simulated (out-of-sample) default correlations by rating for Model II(c) for the final simulation year, 2003. Recall that for the pooled version of Model II(c) the pair-wise correlation of defaults is the same for all firms and is the same as in the Vasicek model (i.e. $\hat{\rho}_{i,j} = \hat{\rho} \ \forall i, j$ where $\hat{\rho}$ is given by equation (76)). Pair-wise correlations of simulated (out-of-sample) defaults are obtained as described in Section 9.4.1 using equation (77).

# Table 8
## The Impact of Parameter Heterogeneity on the Loss Distribution

**Panel A**  The Impact on Expected Losses (EL) of Allowing for Fixed Effects

| Year | Simulated EL | |
|------|--------------|--------------|
|      | **Vasicek**  | **Fixed Effect** |
| **1998** | 1.23% | 1.72% |
| **1999** | 1.60% | 2.17% |
| **2000** | 2.10% | 2.67% |
| **2001** | 2.28% | 2.93% |
| **2002** | 2.74% | 3.28% |
| **2003** | 3.26% | 3.65% |

**Panel B**  The Impact on Unexpected Losses (UL) of Allowing for Fixed Effects

| Year | Simulated UL | |
|------|--------------|--------------|
|      | **Vasicek**  | **Fixed Effect ($\sigma^2$)** |
| **1998** | 1.47% | 1.39% |
| **1999** | 1.94% | 1.84% |
| **2000** | 2.22% | 2.10% |
| **2001** | 1.75% | 1.68% |
| **2002** | 1.98% | 1.90% |
| **2003** | 2.48% | 2.40% |

**Note:** The results in Panel A compare Expected Losses (EL) for the U.S. & Japan Pooled version of Model I (Vasicek) with EL allowing for firm fixed effects, $\alpha_i$, where we have deliberately not equalized EL with Model I. The results in Panel B compare Unexpected Losses (UL) for the U.S. & Japan Pooled version of the Vasicek model (Model I) with UL for Model II(a) - Fixed Effect. In Panel B, we have equalized EL across the two models for each year. The simulation procedure is discussed in Appendix B. All simulations are carried out using 200,000 replications.

# Table 9
## Out-of-Sample Simulated Annual Losses Based on 10-Year Rolling Return Regressions

200,000 replications

| | Simulation Year | Using Sample | | Model Specifications | Default Thresholds | US & Japan Pooled | | | Value-at-Risk | | | US & Japan Modeled Separately | | | Value-at-Risk | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% |
| **Panel A** | 1998 | 1988-1997 | I | Vasicek | $\lambda_c$ / $DD_c$ | 1.47% | 3.1 | 19.5 | 7.00% | 8.48% | 12.05% | 1.29% | 2.6 | 15.5 | 6.10% | 7.22% | 10.14% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 1.39% | 2.8 | 16.9 | 6.57% | 7.86% | 11.14% | 1.23% | 2.5 | 14.2 | 5.80% | 6.82% | 9.51% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 0.80% | 1.1 | 5.1 | 3.67% | 4.08% | 5.11% | 0.77% | 1.1 | 6.9 | 3.47% | 3.82% | 4.75% |
| | $\hat{\pi} =$ | 1.23% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.07% | 1.3 | 5.8 | 4.57% | 5.19% | 6.72% | 1.01% | 1.1 | 5.1 | 4.30% | 4.81% | 6.17% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 0.84% | 0.7 | 3.9 | 3.55% | 3.89% | 4.81% | 0.83% | 0.7 | 4.0 | 3.48% | 3.84% | 4.74% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 0.86% | 1.0 | 5.9 | 3.73% | 4.19% | 5.56% | 0.88% | 1.0 | 5.7 | 3.82% | 4.33% | 5.75% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 0.88% | 1.0 | 6.1 | 3.81% | 4.32% | 5.58% | 0.90% | 1.1 | 5.9 | 3.89% | 4.42% | 5.81% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.08% | 1.6 | 8.8 | 4.82% | 5.62% | 7.69% | 1.07% | 1.5 | 7.8 | 4.74% | 5.47% | 7.30% |
| **Panel B** | 1999 | 1989-1998 | I | Vasicek | $\lambda_c$ / $DD_c$ | 1.94% | 3.0 | 18.1 | 9.30% | 11.13% | 15.91% | 1.73% | 2.8 | 17.9 | 8.25% | 9.77% | 14.04% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 1.84% | 2.8 | 16.3 | 8.76% | 10.43% | 14.75% | 1.65% | 2.7 | 16.3 | 7.86% | 9.31% | 13.12% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 1.02% | 1.3 | 6.2 | 4.84% | 5.43% | 6.86% | 0.97% | 1.7 | 13.9 | 4.55% | 5.15% | 7.02% |
| | $\hat{\pi} =$ | 1.60% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.28% | 1.4 | 6.7 | 5.80% | 6.62% | 8.52% | 1.21% | 1.4 | 6.7 | 5.50% | 6.23% | 7.95% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 0.95% | 0.8 | 4.4 | 4.36% | 4.81% | 5.81% | 0.93% | 0.9 | 5.7 | 4.29% | 4.76% | 5.87% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 0.95% | 1.3 | 8.2 | 4.53% | 5.16% | 6.93% | 0.98% | 1.4 | 8.7 | 4.67% | 5.35% | 7.06% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 0.99% | 1.2 | 7.4 | 4.64% | 5.31% | 7.06% | 1.04% | 1.5 | 8.9 | 4.88% | 5.59% | 7.67% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.20% | 1.7 | 9.3 | 5.69% | 6.58% | 8.88% | 1.20% | 1.8 | 10.4 | 5.72% | 6.70% | 8.95% |
| **Panel C** | 2000 | 1990-1999 | I | Vasicek | $\lambda_c$ / $DD_c$ | 2.22% | 2.6 | 14.7 | 10.65% | 12.59% | 17.34% | 2.06% | 2.4 | 12.7 | 9.78% | 11.61% | 15.87% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 2.10% | 2.4 | 13.1 | 10.03% | 11.81% | 16.17% | 1.97% | 2.3 | 11.7 | 9.34% | 10.97% | 14.87% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 1.29% | 1.4 | 6.9 | 6.29% | 7.06% | 8.99% | 1.21% | 1.3 | 7.9 | 5.85% | 6.59% | 8.69% |
| | $\hat{\pi} =$ | 2.10% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.43% | 1.2 | 5.5 | 6.56% | 7.35% | 9.18% | 1.39% | 1.1 | 5.3 | 6.37% | 7.17% | 8.91% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 1.08% | 0.6 | 3.9 | 5.07% | 5.54% | 6.63% | 1.07% | 0.7 | 4.1 | 5.08% | 5.56% | 6.75% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 1.05% | 1.0 | 6.3 | 5.23% | 5.85% | 7.54% | 1.12% | 1.1 | 6.1 | 5.51% | 6.21% | 7.99% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 1.09% | 1.0 | 5.7 | 5.33% | 5.94% | 7.56% | 1.19% | 1.1 | 6.2 | 5.74% | 6.45% | 8.25% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.39% | 1.5 | 7.7 | 6.68% | 7.69% | 10.16% | 1.40% | 1.6 | 8.4 | 6.77% | 7.79% | 10.46% |

**Table 9 (continued)**
# Out-of-Sample Simulated Annual Losses Based on 10-Year Rolling Return Regressions

200,000 replications

| | Simulation Year | Using Sample | | Model Specifications | Default Thresholds | US & Japan Pooled | | | Value-at-Risk | | | US & Japan Modeled Separately | | | Value-at-Risk | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% |
| **Panel D** | 2001 | 1991-2000 | I | Vasicek | $\lambda_c$ / $DD_c$ | 1.75% | 1.8 | 8.6 | 8.39% | 9.66% | 12.72% | 1.72% | 1.7 | 8.2 | 8.20% | 9.38% | 12.04% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 1.68% | 1.7 | 7.8 | 8.00% | 9.22% | 12.05% | 1.65% | 1.6 | 7.4 | 7.81% | 8.91% | 11.31% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 1.08% | 0.9 | 4.3 | 5.46% | 5.96% | 7.19% | 1.08% | 0.8 | 4.2 | 5.42% | 5.88% | 7.01% |
| | $\hat{\pi} =$ | 2.28% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.22% | 0.7 | 3.8 | 5.71% | 6.23% | 7.38% | 1.23% | 0.7 | 3.9 | 5.71% | 6.23% | 7.40% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 1.00% | 0.4 | 3.1 | 4.84% | 5.15% | 5.86% | 1.02% | 0.4 | 3.2 | 4.92% | 5.28% | 5.99% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 1.13% | 0.8 | 4.4 | 5.54% | 6.09% | 7.46% | 1.20% | 0.8 | 4.5 | 5.76% | 6.31% | 7.69% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 1.18% | 0.7 | 4.2 | 5.64% | 6.20% | 7.57% | 1.27% | 0.8 | 4.4 | 5.96% | 6.58% | 8.06% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.44% | 1.1 | 5.4 | 6.74% | 7.55% | 9.49% | 1.44% | 1.1 | 5.5 | 6.75% | 7.57% | 9.66% |
| **Panel E** | 2002 | 1992-2001 | I | Vasicek | $\lambda_c$ / $DD_c$ | 1.98% | 1.6 | 7.3 | 9.48% | 10.74% | 13.71% | 1.96% | 1.6 | 7.5 | 9.45% | 10.62% | 13.65% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 1.90% | 1.6 | 6.9 | 9.15% | 10.32% | 12.97% | 1.90% | 1.6 | 7.0 | 9.17% | 10.28% | 13.06% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 1.23% | 0.9 | 4.5 | 6.44% | 7.05% | 8.37% | 1.23% | 0.9 | 4.8 | 6.40% | 6.98% | 8.48% |
| | $\hat{\pi} =$ | 2.74% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.34% | 0.5 | 3.4 | 6.32% | 6.80% | 7.91% | 1.35% | 0.6 | 3.5 | 6.39% | 6.91% | 8.07% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 1.10% | 0.3 | 2.9 | 5.43% | 5.76% | 6.47% | 1.11% | 0.3 | 3.1 | 5.52% | 5.86% | 6.66% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 1.16% | 0.5 | 3.7 | 5.87% | 6.38% | 7.55% | 1.22% | 0.6 | 3.9 | 6.08% | 6.61% | 7.89% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 1.19% | 0.6 | 3.8 | 5.98% | 6.50% | 7.73% | 1.26% | 0.7 | 4.1 | 6.25% | 6.82% | 8.23% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.42% | 0.9 | 4.8 | 6.95% | 7.68% | 9.43% | 1.44% | 1.0 | 5.2 | 7.03% | 7.81% | 9.86% |
| **Panel F** | 2003 | 1993-2002 | I | Vasicek | $\lambda_c$ / $DD_c$ | 2.48% | 1.8 | 8.0 | 11.92% | 13.59% | 17.47% | 2.56% | 1.8 | 8.2 | 12.20% | 13.89% | 17.81% |
| | | | II(a) | Fixed Effect ($\sigma^2$) | $\lambda_c$ / $DD_{c,i}$ | 2.40% | 1.7 | 7.5 | 11.54% | 13.06% | 16.84% | 2.48% | 1.7 | 7.8 | 11.84% | 13.46% | 17.07% |
| | | | II(b) | Fixed Effect ($\sigma^2_i$) | $\lambda_c$ / $DD_{c,i}$ | 1.53% | 1.1 | 5.3 | 8.02% | 8.84% | 10.73% | 1.60% | 1.2 | 5.8 | 8.33% | 9.22% | 11.37% |
| | $\hat{\pi} =$ | 3.26% | II(c) | Rating ($\sigma^2$) | $\lambda_i$ / $DD_R$ | 1.51% | 0.6 | 3.7 | 7.46% | 8.06% | 9.46% | 1.56% | 0.7 | 3.9 | 7.65% | 8.28% | 9.86% |
| | | | II(d) | Fixed Effect ($\sigma^2_i$) | $\lambda_i$ / $DD_R$ | 1.16% | 0.3 | 3.2 | 6.24% | 6.64% | 7.53% | 1.22% | 0.4 | 3.5 | 6.46% | 6.91% | 7.95% |
| | | | III(a) | CAPM | $\lambda_i$ / $DD_R$ | 1.27% | 0.8 | 4.7 | 6.98% | 7.64% | 9.21% | 1.27% | 0.8 | 5.0 | 7.00% | 7.66% | 9.30% |
| | | | III(b) | Sector CAPM | $\lambda_i$ / $DD_R$ | 1.28% | 0.8 | 4.7 | 7.01% | 7.64% | 9.20% | 1.30% | 0.9 | 5.2 | 7.15% | 7.83% | 9.60% |
| | | | IV | PCA | $\lambda_i$ / $DD_R$ | 1.51% | 1.1 | 5.9 | 7.96% | 8.93% | 11.15% | 1.49% | 1.2 | 6.2 | 7.91% | 8.85% | 11.11% |

**Note:** This table presents results for simulated out-of-sample annual loss distributions (4-quarter ahead loss distributions). The table presents simulation results for the "pooled" country models in the first set of columns and the "separate" country models in the second set of columns. Model specifications, including both the return specification and determination of default threshold λ, are discussed in Section 9.2 (see Table 3 for more detail on the model specifications). The simulation routine is discussed in Appendix B. Simulation are carried out using 200,000 replications. All models for each year 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), the 3rd and 4th moments of the loss distributions, as well as the 99.0%, 99.5% and 99.9% quantiles of the distribution (denoted Value-at-Risk).

# Table 10
## Comparison of Alternative Methods of Identification of Default Thresholds (2003)

200,000 replications

**Panel A: Same Distance to Default by Rating: $DD_R$**

| Simulation Year | Using Sample | Model Specifications | | US & Japan Pooled | | | | Value-at-Risk | | | US & Japan Modeled Separately | | | | Value-at-Risk | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% |
| **2003** | **1993-2002** | I | Vasicek | 2.48% | 1.8 | 8.0 | 11.92% | 13.59% | 17.47% | | 2.56% | 1.8 | 8.2 | 12.20% | 13.89% | 17.81% |
| | | II(c) | Rating ($\sigma^2$) | 1.51% | 0.6 | 3.7 | 7.46% | 8.06% | 9.46% | | 1.56% | 0.7 | 3.9 | 7.65% | 8.28% | 9.86% |
| $\hat{\pi} =$ | **3.26%** | II(d) | Fixed Effect ($\sigma^2_i$) | 1.16% | 0.3 | 3.2 | 6.24% | 6.64% | 7.53% | | 1.22% | 0.4 | 3.5 | 6.46% | 6.91% | 7.95% |
| | | III(a) | CAPM | 1.27% | 0.8 | 4.7 | 6.98% | 7.64% | 9.21% | | 1.27% | 0.8 | 5.0 | 7.00% | 7.66% | 9.30% |
| | | III(b) | Sector CAPM | 1.28% | 0.8 | 4.7 | 7.01% | 7.64% | 9.20% | | 1.30% | 0.9 | 5.2 | 7.15% | 7.83% | 9.60% |
| | | IV | PCA | 1.51% | 1.1 | 5.9 | 7.96% | 8.93% | 11.15% | | 1.49% | 1.2 | 6.2 | 7.91% | 8.85% | 11.11% |

**Panel B: Same Default Threshold by Rating: $\lambda_R$**

| Simulation Year | Using Sample | Model Specifications | | US & Japan Pooled | | | | Value-at-Risk | | | US & Japan Modeled Separately | | | | Value-at-Risk | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% | | UL | Skew. | Kurt. | 99.0% | 99.5% | 99.9% |
| **2003** | **1993-2002** | I | Vasicek | 2.48% | 1.8 | 8.0 | 11.92% | 13.59% | 17.47% | | 2.56% | 1.8 | 8.2 | 12.20% | 13.89% | 17.81% |
| | | II(c) | Rating ($\sigma^2$) | 1.51% | 0.6 | 3.7 | 7.46% | 8.06% | 9.46% | | 1.56% | 0.7 | 3.9 | 7.65% | 8.28% | 9.86% |
| $\hat{\pi} =$ | **3.26%** | II(d) | Fixed Effect ($\sigma^2_i$) | 1.08% | 0.2 | 2.9 | 5.90% | 6.22% | 6.85% | | 1.12% | 0.2 | 3.0 | 6.04% | 6.37% | 7.09% |
| | | III(a) | CAPM | 1.23% | 0.7 | 4.1 | 6.76% | 7.36% | 8.63% | | 1.21% | 0.6 | 4.1 | 6.64% | 7.19% | 8.53% |
| | | III(b) | Sector CAPM | 1.25% | 0.7 | 4.2 | 6.81% | 7.42% | 8.76% | | 1.24% | 0.7 | 4.4 | 6.83% | 7.43% | 8.86% |
| | | IV | PCA | 1.42% | 0.8 | 4.6 | 7.44% | 8.18% | 9.83% | | 1.39% | 0.9 | 4.8 | 7.39% | 8.13% | 9.82% |

**Note:** This table compares the two identification methods for default thresholds discussed in Section 3 for the final simulation year, 2003. Panel A repeats the five models from Table 9 that are calibrated according to the identifying restriction (11), same distance-to-default by rating. Panel B presents results for these models using identifying restriction (10), namely that all firms with a given rating have the same default threshold, $\lambda_R$. For the "separate" country models, we only impose that all firms in a given country with the same rating have the same threshold $\lambda_{R,c}$. Thus, there are 7 thresholds for the "pooled" models and 14 thresholds for the "separate" country models. The simulation routine is discussed in Appendix B. Simulations are carried out using 200,000 replications. All models for each year are calibrated to have the same Expected Loss given by $\hat{\pi}$. For each simulation, the table reports the standard deviation of losses (denoted Unexpected Losses - UL), the 3rd and 4th moments of the loss distributions, as well as the 99.0%, 99.5% and 99.9% quantiles of the distribution (denoted Value-at-Risk).

# Table 11

## Comparison of $\overline{\hat{\rho}_{i,j}}$ and $\overline{\hat{\rho}_{i,j}^*}$ for Models Using Same Distance to Default Identification (2003)

## Based on U.S. and Japan Pooled Return Specifications

| Simulation Year | Using Sample | Model Specifications | | Parameter Restrictions | $\overline{\hat{\rho}_{i,j}}$ | $\overline{\hat{\rho}_{i,j}^*}$ | Simulated UL |
|---|---|---|---|---|---|---|---|
| **2003** | **1993-2002** | I | Vasicek | $\hat{\pi}_i = \hat{\pi} \ \forall i ; \hat{\rho}_{i,j} = \hat{\rho}, \hat{\rho}_{i,j}^* = \hat{\rho}^* \ \forall i, j$ | 9.20% | 1.80% | 2.48% |
| | | II(c) | Rating ($\sigma^2$) | $\hat{\pi}_i = \hat{\pi}_R \ \forall i \in R ; \hat{\rho}_{i,j} = \hat{\rho} \ \forall i, j ; \hat{\rho}_{i,j}^* = \hat{\rho}_{R,R'}^* \ \forall i \in R, j \in R'$ | 9.20% | 0.26% | 1.51% |
| $\hat{\pi} =$ | **3.26%** | II(d) | Fixed Effects ($\sigma_i^2$) | $\hat{\pi}_i = \hat{\pi}_R \ \forall i \in R ;$ | 9.06% | 0.18% | 1.16% |
| | | III(a) | CAPM | $\hat{\pi}_i = \hat{\pi}_R \ \forall i \in R ;$ | 10.09% | 0.32% | 1.27% |
| | | III(b) | Sector/ CAPM | $\hat{\pi}_i = \hat{\pi}_R \ \forall i \in R ;$ | 10.35% | 0.33% | 1.28% |
| | | IV | PCA | $\hat{\pi}_i = \hat{\pi}_R \ \forall i \in R ;$ | 14.77% | 0.80% | 1.51% |

**Note:** This table compares the average pair-wise correlation of simulated (out-of-sample) returns, $\overline{\hat{\rho}_{i,j}}$ , and the average pair-wise correlation of simulated (out-of-sample) defaults, $\overline{\hat{\rho}_{i,j}^*}$ , across models for the final simulation year, 2003.  These results are only presented for the 6 models that use the same distance-to-default identifying restriction, (11).  Pair-wise correlations of simulated (out-of-sample) returns, $\hat{\rho}_{i,j}$ , are given by equation (18).  Pair-wise correlations of simulated (out-of-sample) defaults, $\hat{\rho}_{i,j}^*$ ,are obtained using equation (77) as described in Section 9.4.1.

**Figure 1**

**Relationship between Default Correlation ($\rho^*$) and Return Correlation ($\rho$) for different values**

**of the degrees of freedom of the Student t-distribution ($\upsilon$) and default probability ($\pi$)**
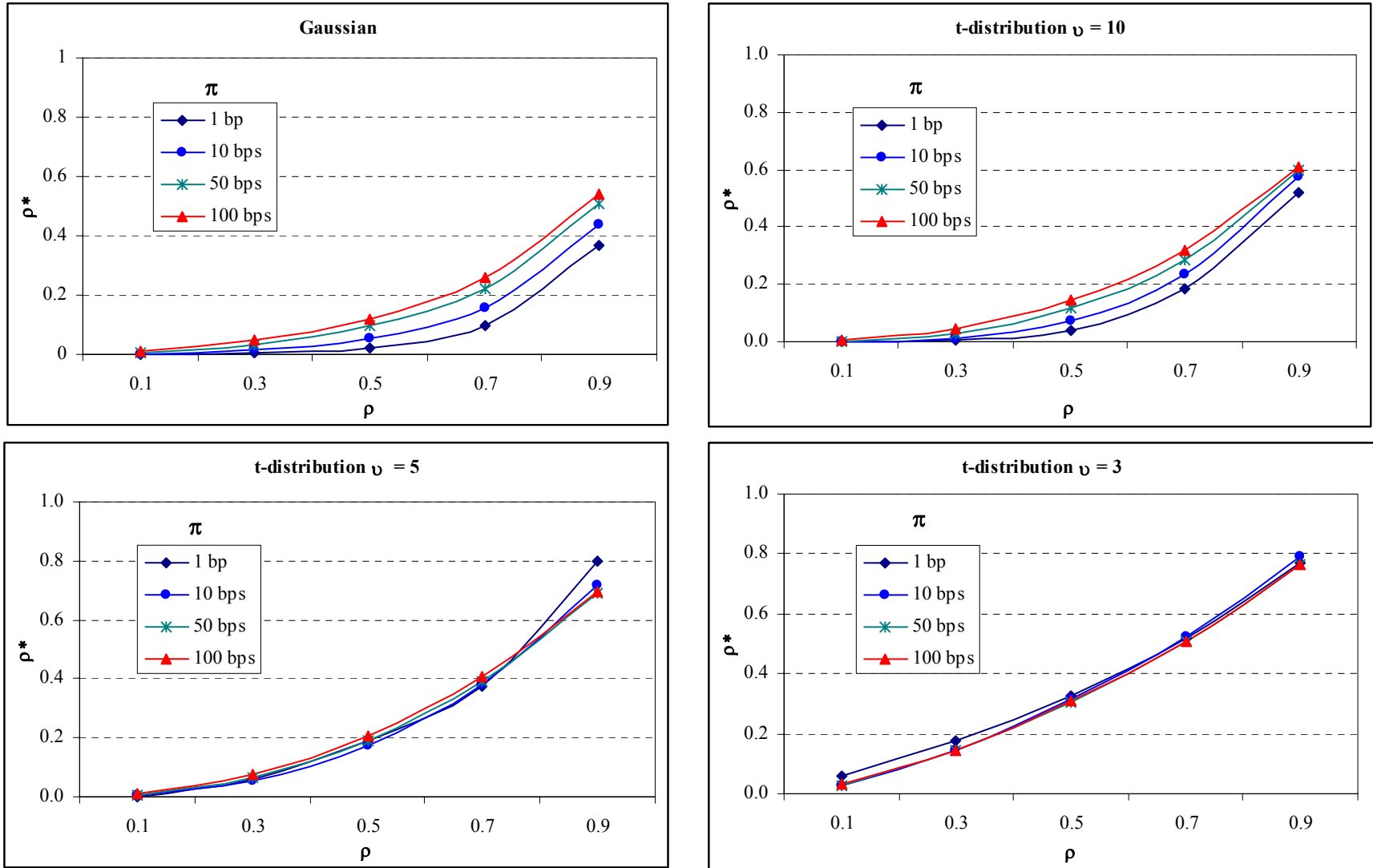
**Note:** $\rho^* = \rho^*(\pi, \rho, \upsilon)$ is calculated using equation (37) for the Student t-distribution and using equation (33) for the Gaussian case. The expectations in equations (33) and (37) are evaluated using 1 million draws of $f_{t+1}$.

59

**Figure 2**
**Kernel Density Estimates of Estimated Betas**
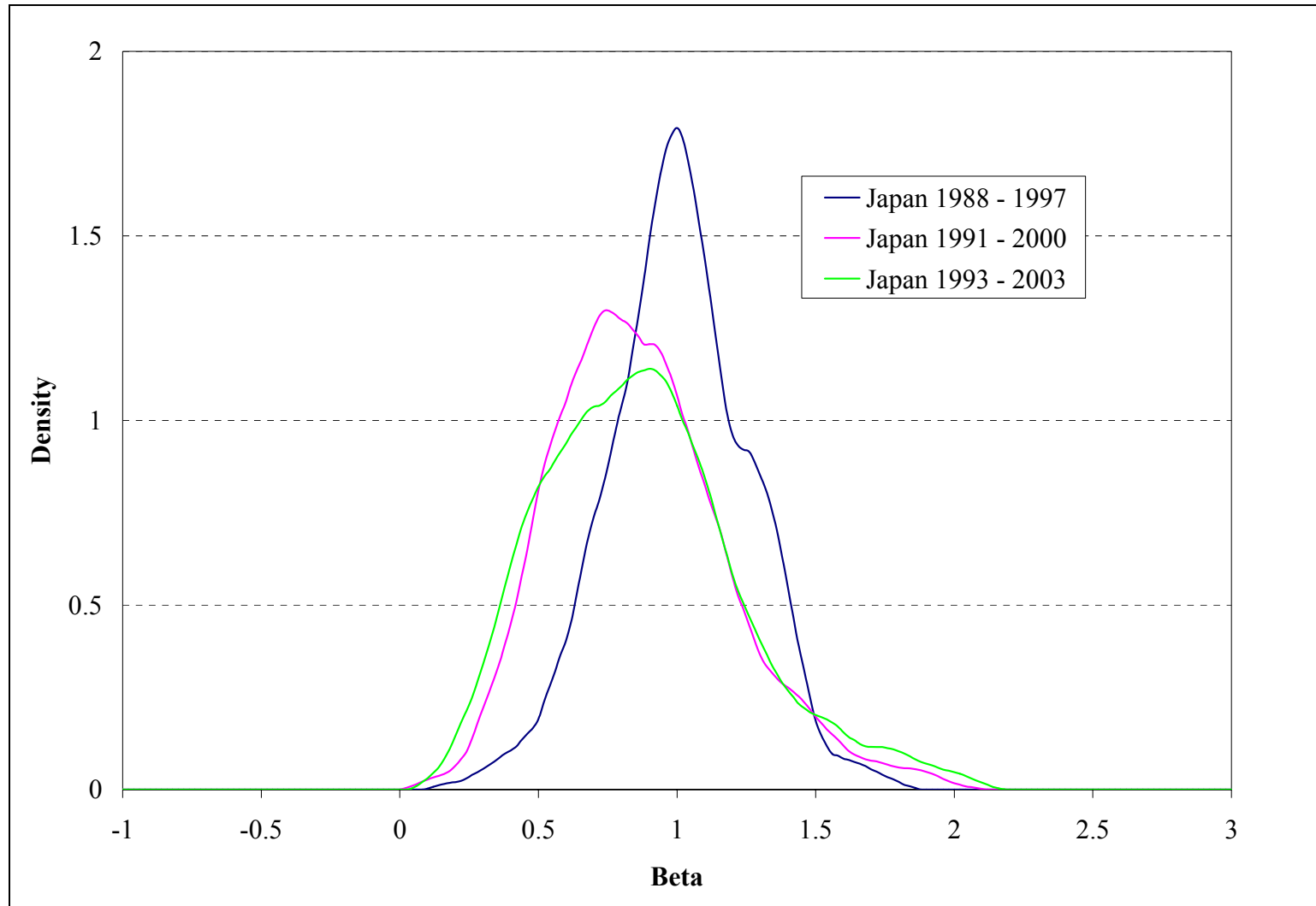**(U.S. versus Japan, 1988-1997)**
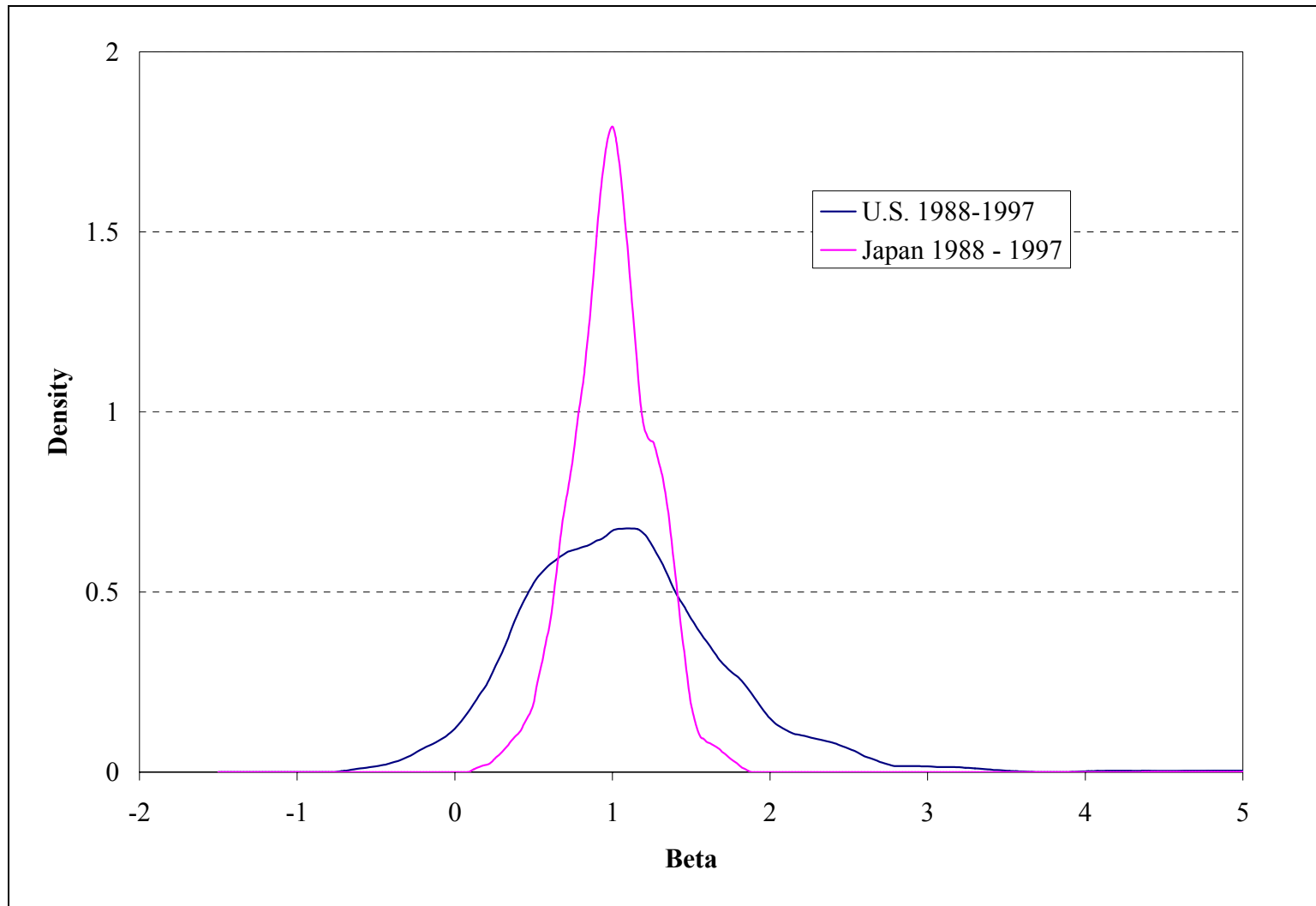


**Note:** Estimated betas are from Model III(a) – CAPM where U.S. & Japan are modeled separately. All densities are estimated with an Epanechnikov kernel using Silverman's (1986) optimal bandwidth.

**Figure 3a**
**Kernel Density Estimates of Estimated Betas**
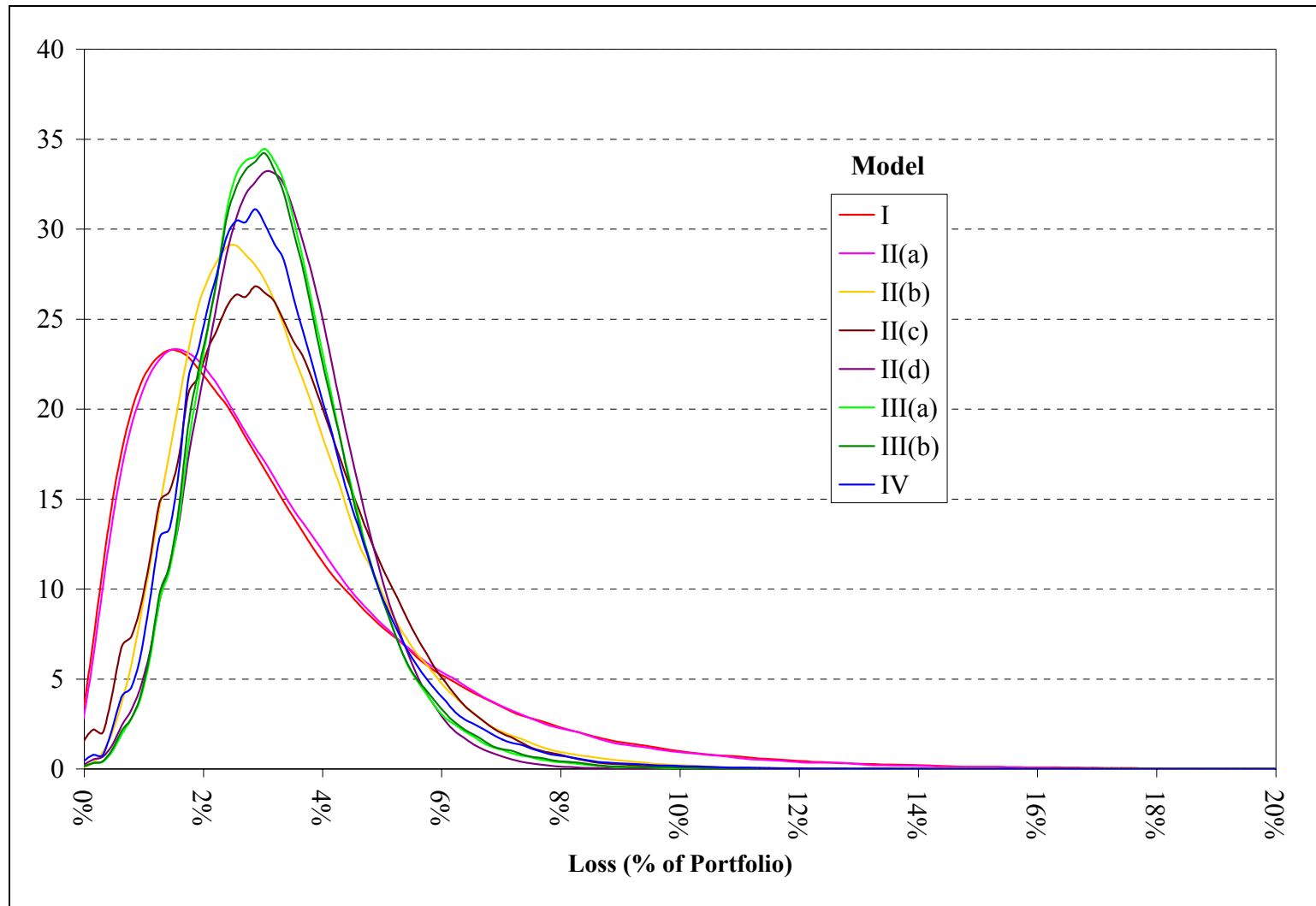**(U.S. Betas over Time)**
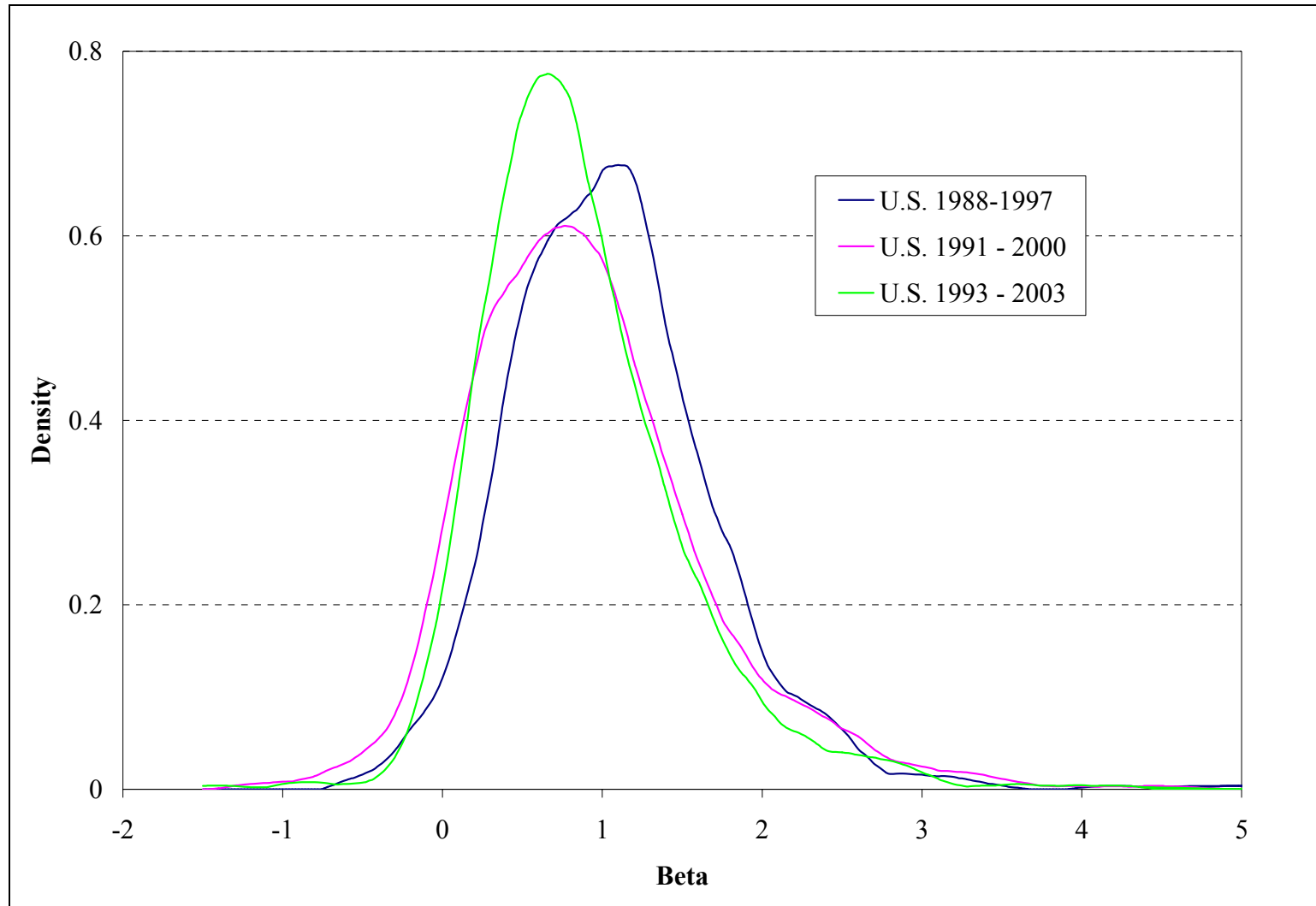


**Note:** Estimated betas are from Model III(a) – CAPM where U.S. & Japan are modeled separately. All densities are estimated with an Epanechnikov kernel using Silverman's (1986) optimal bandwidth.

**Figure 3b**
**Kernel Density Estimates of Estimated Betas**
**(Japanese Betas over Time)**



**Note:** Estimated betas are from Model III(a) – CAPM where U.S. & Japan are modeled separately. All densities are estimated with an Epanechnikov kernel using Silverman's (1986) optimal bandwidth.

# Figure 4
## Alternative Simulated Loss Densities
## (U.S. and Japan Modeled Separately, 2003)



**Note:** All densities are estimated with an Epanechnikov kernel using Silverman's (1986) optimal bandwidth.