Modelling and Calibration Errors in Measures of Portfolio Credit Risk*

Nikola Tarashev†
Haibin Zhu‡

March 2007

(Preliminary. Please do not circulate without permission.)

Abstract

This paper develops an empirical procedure for analysing the impact of model misspecification and calibration errors on measures of portfolio credit risk. On the basis of simulated credit portfolios with realistic characteristics, this procedure reveals that violations of key assumptions of the well-known Asymptotic Single-Risk Factor (ASRF) model are virtually inconsequential, especially for large well-diversified portfolios. By contrast, errors in the calibrated interdependence of credit risk across exposures, which are driven by plausible small-sample estimation errors or popular rule-of-thumb values of asset return correlations, can lead to significant inaccuracies in measures of portfolio credit risk. Similar inaccuracies arise under erroneous, albeit standard, assumptions regarding the tail distribution of asset returns.

JEL Classification Numbers: G21, G28, G13, C15
Keywords: Portfolio credit risk; Value at risk; Bank capital; Credit risk concentration; Estimation errors; Distributional assumptions

---

*The views expressed in this paper are solely those of the authors and do not necessarily represent those of the Bank for International Settlements. We would like to thank Klaus Duellmann, Kostas Tsatsaronis and participants at a BIS workshop for useful comments. Marcus Jellinghaus provided valuable help with the data.

†Nikola Tarashev: Research and Policy Analysis, Bank for International Settlements, Basel, Switzerland. Tel.: 41-61-280-9213. Fax: 41-61-280-9100. E-mail: nikola.tarashev@bis.org.

‡Haibin Zhu: Research and Policy Analysis, Bank for International Settlements, Basel, Switzerland. Tel.: 41-61-280-9164. Fax: 41-61-280-9100. E-mail: nikola.tarashev@bis.org.

1 Introduction

The assessment of portfolio credit risk has attracted much attention in recent years. One reason is that participants in the increasingly popular market for structured finance products[1] rely heavily on estimates of the inter-dependence of credit risk across various exposures. Such estimates are also of principal interest to financial regulators who have to enforce new revisions to capital standards in the banking and insurance industries and, thus, ensure that regulatory capital is closely aligned with credit risk. By extension, the validation of portfolio credit risk models is of great interest to both market participants and regulators.

This paper investigates the well-known Asymptotic Single-Risk Factor (ASRF) model of portfolio credit risk. The popularity of this model stems from its implication – derived rigorously in Gordy (2003) – that the capital buffer allocated to cover unexpected losses for a portfolio, ie the credit value-at-risk (VAR) net of expected default losses, is the simple sum of capital buffers set at the level of individual exposures. This implication – also known as *portfolio invariance* of capital buffers – has been often interpreted as alleviating the data requirements and computational burden on users of the model.[2] The reason for this interpretation is that portfolio invariance implies that portfolio-level capital buffers can be calculated solely on the basis of exposure-specific parameters: ie individual probabilities of default (PD), losses-given-default (LGD) and dependence on the common factor.

Its popularity notwithstanding, the "portfolio invariance" implication of the ASRF model hinges on two strong assumptions that have been criticised as sources of *misspecification errors*. Namely, the model assumes that the systematic component of credit risk is governed by a single common factor and that the portfolio is so finely grained that all idiosyncratic risk is diversified away. Violations of the "single-factor" and "perfect granularity" assumptions can affect assessments of portfolio credit risk and, consequently, lead to erroneous estimates of capital buffers.

Moreover, an application of the ASRF model may be quite challenging, even in the absence of misspecification errors. An important reason for this is that the "portfolio invariance" implication does not remove fully the necessity to adopt a global approach in calibrating the model. In particular, estimates of exposure-specific dependence on a common factor, which are required for an application of the ASRF model, hinge on an estimate

[1]These products include collateralized debt obligations (CDOs), nth-to-default credit default swap (CDS) and CDS indices.

[2]The portfolio invariance property has also been incorporated in the internal-ratings-based (IRB) approach of the Basel II framework (BCBS, 2005).

of the common factor itself, which is a market-wide variable.[3] Noise in such estimates would lead to a *flawed calibration* of the ASRF model, which would be another source of errors in measured portfolio credit risk.

A contribution of this paper is to propose a *unified* method for quantifying the importance of model misspecification and calibration errors in assessments of portfolio credit risk. In order to implement this method, we rely on a large data set that comprises Moody's KMV estimates of PDs and pairwise asset return correlations for nearly 11,000 non-financial corporates worldwide. Treating these estimates as actual credit risk parameters and constructing hypothetical portfolios that match the industrial-sector concentration of typical portfolios of US wholesale banks, we derive the "true" probability distribution of default losses. We summarise this probability distribution, which does not rely on the ASRF model, into a credit VAR net of expected losses. This summary measure is equivalent to a "target" capital measure necessary to cover default losses with a desired probability.[4] Then, our method dissects the difference between the target capital for a given portfolio and a "shortcut" capital, which is calculated for the same portfolio on the basis of the ASRF model and a rule-of-thumb calibration of exposure-specific dependences on a common risk factor.

We decompose the difference between the target and shortcut capital measures into four non-overlapping and exhaustive components. Two of these components relate to two sources of misspecification of the ASRF model and are attributed to a "multi-factor" effect and a "granularity" effect. The other two components relate to two types of errors in the calibration of the interdependence of credit risk across exposures.[5] Specifically, these errors arise from an overall bias in the calibrated correlations of firms' asset returns ("correlation level" effect) and from noise in the calibrated dispersion of these correlations across pairs of firms ("correlation dispersion" effect).

Another contribution of this paper is that it provides two additional perspectives on flaws in the calibration of the ASRF model. First, we derive plausible ranges of calibration errors that arise not from the adoption of "rule-of-thumb" values but as a result of estimating asset return correlations on the basis of finite data on asset returns. We then convert these

[3]The IRB capital formula of Basel II avoids the necessity of a global approach by postulating that firm-specific dependence on the single common factor is determined fully by the level of the corresponding PD.

[4]In this paper, we use the terms "assessment of credit risk" and "capital measure" interchangeably. Importantly, our capital measures do not correspond to "regulatory capital", which reflects considerations of bank supervisors, or to "economic capital", which reflects additional strategic and business objectives of financial firms.

[5]In order to sharpen the analysis, we do not analyse the implications of errors in the estimates of PDs and LGDs.

small-sample errors into deviations from a desired capital buffer. Second, we examine the importance of an erroneous calibration of the overall distribution of firms' asset returns. Such a calibration would affect the measured interdependence of credit risk across exposures over and above the impact of errors in estimated asset return correlations.

Our main conclusion is that errors in the practical implementation, as opposed to the specification, of the ASRF model are the main causes of flaws in assessments of portfolio credit risk. Specifically, the misspecification-driven multi-factor effect would lead a user of the model to miss target capital buffers by only 1% of their level. This is because a one-factor approximation fits well the correlation structure of asset returns in our data. Similarly, the granularity effect in realistically large portfolios causes calculated capital to deviate from the target level by the negligible 4%. By contrast, capital measures are significantly more sensitive to plausible miscalibration of the ASRF model. For instance, missing the empirical dispersion of asset return correlations across pairs of firms would cause a 12% deviation from target capital. Furthermore, plausible small-sample errors (arising when users of the model have 10 years of monthly asset returns data) could affect substantially asset return correlation estimates, translating into a capital measure that deviates from the target level by 23%. Finally, available data on asset returns suggests that their empirical distribution is at odds with the conventional normality assumption. As a result, adopting this assumption leads to an underestimate of target capital by 11 to 22%.

In comparison to articles in the related literature, this paper covers a wider range of errors in assessments of portfolio credit risk. The related literature has focused mainly on misspecifications of the ASRF model and has proposed ways to both partially correct for them and preserve the tractability of the model. Empirical analyses of violations of the perfect granularity assumption include Martin and Wilde (2002), Vasicek (2002), Emmer and Tasche (2003), and Gordy and Luetkebohmert (2006). For their part, Pykhtin (2004), Duellmann (2006), Garcia Cespedes et al. (2006) and Duellmann and Masschelein (2006) have analysed the validity of the common factor assumption under different degrees of portfolio concentration in a limited number of industrial sectors. In addition, Heitfield et al. (2006) and Duellman et al. (2006) examine both granularity and sector concentration issues in the context of US and European bank portfolios, respectively.[6] Articles that stand apart in the extant related literature are Loeffler (2003) and Morinaga and Shiina (2005), which focus exclusively on estimation-related issues and derive that noise in model parameters can have

[6]The recent working paper by the Basel Committee on Bank Supervision (BCBS, 2006) provides an extensive review of the related literature.

a significant impact on assessments of portfolio credit risk.

The remainder of this paper is organised as follows. Section 2 outlines the ASRF model and the empirical methodology applied to it. Section 3 describes the data and Section 4 reports the empirical results. Finally, Section 5 concludes.

2 Methodology

In this section, we first outline the ASRF model. Then, we discuss how violations of its key assumptions or flawed calibration of its parameters can affect assessments of portfolio credit risk. Finally, we develop an empirical methodology for deriving and comparing alternative sources of errors in such assessments.

2.1 The ASRF model[7]

The ASRF model of portfolio credit risk – introduced by Vasicek (1991) – postulates that an obligor defaults when the value of its assets falls below some threshold. In addition, the model assumes that asset values are driven by a single common factor:

$$V_{iT} = \rho_i \cdot M_T + \sqrt{1-\rho_i^2} \cdot Z_{iT} \tag{1}$$

where: V_{iT} is the value of assets of obligor i at time T; M_T and Z_{iT} denote the underlying values of the common and idiosyncratic factors, respectively; and $\rho_i \in [-1, 1]$ is the obligor-specific loading on the common factor. The common and idiosyncratic factors are independent of each other and can be rescaled, without loss of generality, to random variables with mean 0 and variance 1. Thus, the asset return correlation between borrowers i and j is given by $\rho_i \rho_j$.

The ASRF model delivers a closed-form approximation to the probability distribution of default losses on a portfolio of N exposures. The accuracy of the approximation increases when $N \to \infty$ and the largest exposure weight $\sup_i (w_i) \to 0$. In these limits – ie when the portfolio is perfectly granular – the probability distribution of default losses can be derived as follows. First, let the indicator $\mathcal{I}_{iT}$ equal 1 in the event of a default at time T and 0 otherwise. Conditional on the value of the common factor, the expectation of the indicator

[7] This section provides an intuitive discussion of the ASRF model. For a rigorous analysis of the model, see Gordy (2003).

equals

$$\begin{aligned} E(\mathcal{I}_{iT}|M_T) &= \Pr(V_{iT} < \mathcal{F}^{-1}(PD_{iT})|M_T) \\ &= \Pr(\rho_i \cdot M_T + \sqrt{1-\rho_i^2} \cdot Z_{iT} < \mathcal{F}^{-1}(PD_{iT})|M_T) \\ &= \mathcal{H}\left(\frac{\mathcal{F}^{-1}(PD_{iT}) - \rho_i M_T}{\sqrt{1-\rho_i^2}}\right) \end{aligned}$$

where: PD_{iT} is the probability of default by obligor i at time T; the cumulative distribution function (CDF) of Z_{iT} is denoted by $\mathcal{H}(\cdot)$; the CDF of V_{iT} is $\mathcal{F}(\cdot)$ and implies that the default threshold equals $\mathcal{F}^{-1}(PD_{iT})$.

Second, under perfect granularity, the Law of Large Numbers implies that the conditional total loss on the portfolio, $TL|M$, is deterministic for any value of the common factor M:

$$\begin{aligned} TL|M &= \sum_i w_i \cdot E(\mathcal{I}_i|M) \cdot LGD_i \\ &= \sum w_i \cdot \mathcal{H}\left(\frac{\mathcal{F}^{-1}(PD_i) - \rho_i M}{\sqrt{1-\rho_i^2}}\right) \cdot LGD_i \end{aligned} \tag{2}$$

where LGD_i is the loss-given default of obligor i (assumed here to be known ex ante) and time subscripts have been suppressed.

Finally, by equation (2), the conditional total loss $TL|M$ is a decreasing function of the common factor M and, thus, the unconditional distribution of TL can be derived directly on the basis of the distribution, $\mathcal{G}(\cdot)$, of the common factor. Denoting by TL_α the $(1-\alpha)^{th}$ percentile in the distribution of total losses (ie $\Pr(TL > TL_a) = \alpha$):

$$\begin{aligned} TL_\alpha &= \sum w_i \cdot \mathcal{H}\left(\frac{\mathcal{F}^{-1}(PD_i) - \rho_i \mathcal{G}^{-1}(\alpha)}{\sqrt{1-\rho_i^2}}\right) \cdot LGD_i \\ &= TL|M_\alpha \end{aligned}$$

where M_α is the α^{th} percentile in the distribution of the common factor.

Thus, in order to cover unexpected (ie total minus expected) losses with probability

$(1 - \alpha)$, the capital buffer for the entire portfolio should be set to:

$$\begin{aligned} \kappa &= TL_\alpha - \sum w_i \cdot PD_i \cdot LGD_i \\ &= \sum w_i \cdot LGD_i \cdot [\mathcal{H}\left(\frac{\mathcal{F}^{-1}(PD_i) - \rho_i \mathcal{G}^{-1}(\alpha)}{\sqrt{1-\rho_i^2}}\right) - PD_i] \\ &\equiv \sum w_i \cdot \kappa_i \end{aligned} \tag{3}$$

As implied by this equation, the capital buffer for the portfolio can be set on the basis of exposure-specific parameters. These parameters reflect the weight of a given exposure, its LGD and PD as well as the underlying dependence on the common factor. The flip side of this implication is that each exposure-specific portion of the capital buffer is independent of the rest of the portfolio and, thus, is *portfolio invariant*.

In practice, an implementation of the ASRF model requires that one specify the distribution of the common and idiosyncratic factors of asset returns. It is standard to assume normal distributions, which implies that equation (3) can be rewritten as:

$$\kappa = \sum w_i \cdot LGD_i \cdot [\Phi\left(\frac{\Phi^{-1}(PD_i) - \rho_i \Phi^{-1}(\alpha)}{\sqrt{1-\rho_i^2}}\right) - PD_i] \tag{4}$$

where $\Phi(\cdot)$ is the CDF of a standard normal variable.[8]

2.2 Impact of model misspecification

The portfolio invariance implication of the ASRF model hinges on two key assumptions, ie that the portfolio is of perfect granularity and that there is a single common factor. In this section, we examine at a conceptual level how a violation of either of these assumptions affects capital measures. We dub the two misspecification effects "granularity" and "multi-factor" effects and, in providing specific examples, use the formula in equation 4 and set $\alpha = 0.1\%$.

2.2.1 Granularity effect

The granularity effect arises empirically either because of a limited number of exposures or because of exposure concentration in a small number of borrowers. In either of these cases, idiosyncratic risk is not fully diversified away. Therefore, the existence of a granularity effect

[8] Equation (4) underpins the regulatory capital formula in the IRB approach of Basel II.

implies that capital measures based on the ASRF model are lower than actual unexpected losses.

The top-left panel in Figure 1 provides an illustrative example of the granularity effect. In this example, the target capital level for a homogenous portfolio is computed as a function of the number of exposures (solid line). In addition, the figure also plots the capital measure implied by the ASRF model (dotted line), which differs from the target one only in that it assumes an infinite number of exposures. The difference between the dotted and solid lines equals the magnitude of the granularity effect. As expected, the granularity effect is always negative[9] and decreases when the number of exposures increases.

2.2.2 Multi-factor effect

If the impact of various macroeconomic and industry-specific conditions on credit risk is best summarised by *multiple* mutually independent (and potentially unobservable) common factors of firms' assets, then the single-factor assumption of the ASRF model would be violated. Such a violation leads to what we call a multi-factor effect, which is conceptually different from a failure to measure the impact of multiple factors on the correlation across obligors. Such a failure is independent of a modelling misspecification and can arise, for example, when higher concentration in a particular industrial sector is not captured in the estimated average correlation. However, even if the average correlation across obligors is measured accurately, an erroneous single-factor assumption ignores the fact that there are multiple sources of default clustering. This leads to an underestimation of the probability of a large number of defaults and, consequently, to an underestimation of target capital.

An illustrative example of the multi-factor effect is provided in the top-right panel of Figure 1. For this example, we construct a homogenous portfolio in which the exposures can be divided into two groups. In addition to idiosyncratic factors, the credit risk of the portfolio is driven by two common factors that are group specific. Concretely, this translates into homogenous within-group pairwise correlations that equal 20% and across-group correlations that equal zero. The solid line plots the target capital measure, which incorporate the multi-factor correlation structure, as a function of the relative weight of exposures in group 1. This measure is lowest when the portfolio is most diversified between the two groups of exposures, which minimises the probability of large default losses. In addition, the figure portrays a dashed line, which portrays an alternative capital measure

[9] Note that the granularity effect should be compared to the *negative* of the granularity adjustment derived in Gordy and Luetkebohmert (2006). Such a comparison is reported in footnote 24 below.

underpinned by a single-factor structure of asset return correlations. This structure matches exactly the true *average* asset return correlation but misses the variability of correlation coefficients in the cross section.

The difference between the dashed and solid lines equals the multi-factor effect. This difference is largest when the two groups enter the portfolio with equal weights. In this case, the role of multiple factors is greatest and, hence, a single-factor structure approximates most poorly the differences among the correlation coefficients of asset returns.

2.3 Impact of calibration errors

Errors in the calibration of the ASRF model can affect assessments of portfolio credit risk over and above any effects of model misspecification. In this paper, we consider errors in the calibration of the interdependence of credit risk across exposures. Such errors can be driven by flawed values of asset return correlations or by an empirically violated assumption regarding the distribution of asset returns.

Errors in asset return correlation estimates may arise for various reasons. One possibility is that a user of the ASRF model is data constrained and relies on rule-of-thumb values, which may simply be correlation estimates for popular credit indices. Such estimates would lead to a discrepancy between target and calculated capital to the extent that the underlying indices were not representative of the user's own portfolio. Alternatively, a user of the model may have data on the assets of the obligors in its portfolio but these data may cover a short time period and, thus, lead to small-sample estimation errors in asset return correlations. Such data limitations are likely to be important in practice because: (i) asset value estimates are typically available at the monthly or quarterly frequency and (ii) supervisory texts require from financial institutions five years of relevant data.[10]

A positive error in the average level of asset return correlations leads to a capital measure that is higher than the target one. This is illustrated in the bottom-left panel of Figure 1 and reflects the intuition that a higher level of asset return correlation inflates the probability that a large number of defaults may occur simultaneously. In the remainder of this paper, the implications of such errors are dubbed the "correlation level" effect.

[10]This observation is likely to hold irrespective of the way a user of the model obtains estimates of asset values. Without relying on a data provider, a financial institution may be able to estimate asset values directly from the balance sheets of its obligors. In addition, the assets of publicly listed obligors can be deduced from their stock market prices. Alternatively, Tarashev and Zhu (2006) derive estimates of asset return correlations on the basis of credit-default-swap (CDS) data.

In addition, there could be errors in the estimated dispersion of asset return correlations across exposure pairs.[11] The effect of such errors on calculated capital can be understood in a stylised example. Suppose that all firms in one portfolio have homogeneous PDs and exhibit homogeneous pairwise asset return correlations. Suppose further that a second portfolio is characterised by the same PDs and average asset return correlation but includes a group of firms that are more likely to default together. The second portfolio, in which pairwise correlations exhibit dispersion, is more likely to experience several simultaneous defaults and, thus, requires higher capital in order to attain solvency with the same probability. This is a particular instance of the "correlation dispersion" effect and is portrayed by the upward slope of the solid line in the bottom-right panel of Figure 1.

This result can be strengthened (dashed line in the same panel) but also weakened or even reversed if PDs vary across firms. To see why, suppose that the strongly correlated firms in the second portfolio are the ones that have the lowest individual PDs. In other words, the firms that are likely to generate multiple defaults are less likely to default. This may lower the probability of default clustering, depressing the target capital level below that for the first portfolio. This is illustrated by the dash-dotted line in the bottom-right panel of Figure 1, which has a negative slope for certain degrees of dispersion in asset return correlations.

Even if asset return correlations were known, a flawed calibration of the distribution of asset returns would still lead to an error in the calibrated interdependence of credit risk across exposures. The literature on *pricing* portfolio credit risk has pointed out that empirical asset return distributions have fatter tails than those imposed by the conventional assumption of normal distributions. To the extent that this fatness of the tails reflects the distribution of the common factor, the probability of default clustering and, thus, the target capital level would be higher than those implied under normality (Hull and White, 2004; Tarashev and Zhu, 2006). This effect would be observationally indistinguishable from the correlation level effect and could be studied by considering more general, eg Student t, distributions of the common and idiosyncratic factors of asset returns.

[11] Note that there are two sources of dispersion in asset return correlations. First, as discussed in Section 2.2.2, correlations may differ across exposure pairs if the assets of different firms are driven by different common factors. Second, there may be dispersion across correlation coefficients if there is a single common factor but the loadings on it differ across firms. This case is examined in the present subsection.

2.4 Evaluating various sources of error

An important contribution of this paper is a unified empirical method for quantifying the impact of several sources of error in model-based assessment of portfolio credit risk. In particular, we construct the difference between target capital measures and shortcut ones, based on the ASRF model and possible erroneous calibration of its parameters. Then, we dissect this difference into four overlapping and exhaustive components, attributing them to the multi-factor, granularity, correlation level and correlation dispersion effects. In order to probe further the likely magnitude of the last two effects, we derive plausible small-sample errors that could affect direct estimates of asset return correlations. Finally, we also examine the extent to which the correlation level effect can be augmented by an erroneous assumption regarding the distribution of asset returns.

The basic empirical method consists of two general steps. In the first step, we construct a large (small) hypothetical portfolio that comprises equal exposures to 1,000 (200) firms.[12] The sectoral composition of this portfolio is designed to be in line with the typical loan portfolio of large US wholesale banks.[13] Given the constraints of such a composition, the portfolio is drawn at random from our sample of firms. Since each draw could be affected by sampling errors, we examine 3,000 different draws for both large and small portfolios.

For a portfolio constructed in the first step, the second step calculates five alternative capital measures, which differ in the underlying assumptions regarding the interdependence of credit risk across exposures. Each of these alternatives employs the same set of PD values, and assumes that LGD equals 45% for all exposures and that asset returns are normally distributed. We order the measures so that each measure differs from a previous one owing to a single assumption.

1. The **target capital** measure incorporates data on asset return correlations. Using these correlations, we conduct Monte Carlo simulations to construct the "true" probability distribution of default losses at the one-year horizon. Then we set the target capital to a level that covers unexpected default losses with a probability of 99.9% (see Appendix A for detail).

2. The second capital measure differs from the target one only owing to a restriction

[12] The distinction between large and small portfolios does not reflect the size of the aggregate exposure but rather different degrees of diversification across individual exposures.

[13] Such a portfolio does not incorporate consumer loans and, thus, may not be representative of all aspects of credit risk.

on the number of common factors governing asset returns. In particular, we adopt a correlation matrix that fits the original one as closely as possible under the constraint that correlation coefficients should be consistent with the presence of only one common factor (see Appendix B). The fitted matrix is then used to derive the one-year probability distribution of joint defaults on the basis of the so-called Gaussian copula method (see Appendix C). This distribution is then mapped into a probability distribution of default losses and, finally, into a capital measure.

3. The third capital measure differs from the second one only in that it assumes that all idiosyncratic risk is diversified away. This assumption allows us to incorporate the fitted one-factor correlation matrix in the ASRF formula (equation 4).

4. The fourth capital measure differs from the third one only in that it is based on the assumption that loading coefficients on the single common factor are the same across exposures. The resulting common correlation coefficient, which is set equal to the average of the pairwise correlations underpinning measure 3., is used as an input to the ASRF formula (equation 4).

5. Finally, the **shortcut** capital measure differs from the fourth one only in that it incorporates alternative, rule-of-thumb, values for the common correlation coefficient.

A dissection of the difference between the target and shortcut capital measures is a simple by-product of this methodology.[14] Specifically, the difference between measures 5 and 1 is the sum of the following four components: (i) the difference between measures 2 and 1, which equals the multi-factor effect; (ii) the difference between measures 3 and 2, which equals the granularity effect; (iii) the difference between measures 4 and 3, which equals the correlation dispersion effect; and (iv) the difference between measures 5 and 4, which equals the correlation level effect.

A possible criticism of this methodology is that the shortcut measure is based on an arbitrarily specified rule-of-thumb correlation coefficient. In order to address this issue, we derive a plausible range of errors in pairwise correlation estimates that is generated by realistic limitations on the size of relevant data. Specifically, we draw time series of asset returns from a joint distribution characterised by constant pairwise correlations equal to the correlation underpinning measure 4. Using the sample correlation matrix of the simulated

[14]Importantly, the method also applies to alternative definitions of *target* and *short-cut* capital, so long as the *true* correlation structure and short-cut correlation estimates chosen by the user are clearly defined.

series, together with a typical value for the probability of default, we derive an estimate of the probability distribution of joint defaults. We then incorporate this probability distribution in the ASRF formula (equation 4) to obtain an "estimated" capital measure.[15] Finally, we quantify the estimation errors in this capital measure, which are driven by small-sample noise in the measured overall level and dispersion of asset return correlations.

Another possible criticism of our basic methodology is that the shortcut capital measure incorporates the arguably unrealistic assumption that all underlying distributions are normal. We relax this assumption by generalising the shortcut measure to accommodate Student t distributions of the common and/or idiosyncractic factors of asset returns. This requires using the general ASRF formula in equation (3) and making two technical adjustments. The first adjustment addresses the fact that the variance of a Student t variable is larger than unity.[16] The second adjustment addresses the fact that the generalised CDF of asset returns, $\mathcal{F}(\cdot)$, does not exist in closed form. In concrete terms, we calculate the default threshold $\mathcal{F}^{-1}(PD_i)$ on the basis of 10 million Monte Carlo simulations.

3 Data description

This section describes the two major blocks of data used in this study: (i) the risk parameter estimates provided by Moody's KMV and (ii) the sectoral distribution of exposures in typical portfolios of US wholesale banks.

3.1 Market-based estimates of risk parameters

Our sample includes the universe of firms covered in July 2006 by two proprietary products of Moody's KMV: the expected default frequency (EDF) model and the global correlation (GCorr) model. An EDF is an estimate of the 1-year physical PD of a publicly traded firm, while the GCorr model delivers an estimate of the physical pairwise asset return correlation between any two publicly traded firms in the dataset. We abstract from financial firms – whose capital structure makes their PDs notoriously difficult to estimate – and work with 10,891 firms.

The sample covers firms with diverse characteristics. Specifically, 5,709 of the firms are headquartered in the United States, 4,383 in Western Europe, and the remaining 799 in the rest of the world. The distribution of the 10,891 firms across industrial sectors is reported

[15] These calculations abstract from the granularity and multi-factor effects.

[16] Specifically a Student t variable with $r > 2$ degrees of freedom has a variance of $\frac{r}{r-2}$.

in the last column in Table 1, with the largest share of firms (about 10.4%) coming from the business service sector. Importantly, only 1,434 (or 13.2%) of the firms have a rating from either S&P or Moody's, which matches the stylised fact that the majority of bank exposures are unrated.

The two proprietary products, EDF and GCorr correlations, are derived within a coherent framework, built on a Merton (1974) type model. This model, largely in line with the ASRF model, postulates that a default occurs when the borrower's asset value falls below a threshold (see Crosbie and Bohn, 2003, for detail). EDFs have been widely used as proxies for physical default probabilities (see Berndt et al., 2005; Longstaff et al., 2005, for example). In turn, GCorr correlations rely on model-implied time series of asset values and a multi-factor structure of asset return correlations. In particular, this model incorporates 120 common factors, including 2 global economic factors, 5 regional economic factors, 7 sector factors, 61 industry-specific factors and 45 country-specific factors (see Das and Ishii, 2001; Crosbie, 2005).

Table 2 and Figure 2 report summary statistics of the Moody's KMV 1-year PD and asset return correlation estimates. PDs have a long right tail and, thus, their median (0.39%) is much lower than the mean (2.67%). In addition, the favourable credit conditions in July 2006 have resulted in 1,217 firms (ie about 11.2% of the total) having the lowest EDF score (0.02%) allowed by the Moody's KMV empirical methodology. For comparison, the upper bound on the Moody's KMV PD estimates (20%) is attained by 643 (or 5.9%) of firms in the sample. For pairwise correlation coefficients, GCorr imposes a lower limit of 0 and an upper limit of 65%. The majority of pairwise correlation coefficients are between 5% and 25%, while the mean stands at 9.24%.[17]

3.2 Constructing hypothetical portfolios

In constructing hypothetical portfolios, we mimic the sectoral distribution of large US wholesale banks, as reported in Heitfield et al. (2006). Specifically, to construct a large portfolio (1000 exposures), we apply the 40 non-financial sector weights reported by that paper (see Table 1). For a small portfolio (200 exposures), we rescale the 10 largest sectoral weights so that they sum up to unity and set all other weights to zero. Within each sector, we

[17]The GCorr correlation estimates are quite in line with correlation estimates reported in other studies. For instance, Lopez (2004) documents an average asset correlation of 12.5% for a large number of US firms and Duellman et al. (2006) estimate a median asset return correlation of 10.1% for European firms.

draw firms at random.[18] All firms in a portfolio receive equal weights and, thus, there is a one-to-one correspondence between the number of firms in a sector and this sector's weight in the portfolio.

The sector and name concentration indices, reported at the bottom of Table 1 are in line with those reported in Heitfield et al. (2006). The sector (name) concentration index of large portfolios studied in this paper is calculated as the sum of squared sectoral (name) weights and equals 0.0432 (0.001), which belongs to the interval $[0.03, 0.045]$ ($[0.000, 0.003]$) reported in Heitfield et al. (2006). For small portfolios the corresponding indices and intervals are 0.1135 (0.005) and [0.035, 0.213] ([0.001, 0.008]), respectively.

4 Empirical results

We implement the empirical methodology described in Section 2.4 and quantify the impact of various sources of error in ASRF-based assessments of portfolio credit risk. Before reporting the empirical results, it is useful to highlight several aspects of the methodology.

First, as far as calibration of the model is concerned, the analysis in this paper focuses exclusively on errors in the values of parameters that relate to the interdependence of credit risk across exposures. Thus, we abstract from potential errors in the calibration of individual PDs and LGDs. Specifically, all capital measures we consider are based on Moody's KMV PD estimates, which we treat as being free of estimation error, and on LGDs set to 45%, which is a rule-of-thumb value in the literature. Considering the impact of noise in PD and LGD values would make it extremely difficult to isolate the correlation level and dispersion effects we focus on. This is because noise in PDs and LGDs would interact with noise in correlation inputs in a highly non-linear fashion.

Second, we make the stylised assumption that portfolios consist of equally weighted exposures. Considering disparate exposure sizes would require considering an additional dimension of portfolio characteristics, as it will no longer be the case that the granularity of a larger portfolio is necessarily finer. In addition, lower granularity that results from higher concentration in a small number of borrowers would also have a bearing on the number of common factors affecting the portfolio and on the overall correlation of risk. This would make it impossible to isolate the granularity effect from the other three effects we consider.

[18] In drawing firms from the sample set, we choose to draw randomly *with* replacement within each industry. If the same firm is drawn twice, the corresponding pairwise correlation is set equal to the average correlation in the industry sector. Drawing randomly *without* replacement does not affect materially the results.

Finally, our analysis treats the correlation matrix provided by Moody's KMV as providing the "true" correlation of asset returns. Of course, this matrix is itself an estimate and is subject to errors. Nevertheless, the Moody's KMV correlation matrix provides a reasonable benchmark to work from. In addition, we have verified that results regarding the *relative* importance of alternative sources of error depend only marginally on the accuracy of the GCorr estimates, even though the *absolute* impact of alternative sources of error does change with the benchmark correlation level.

4.1 Various errors in shortcut capital measures

To study various sources of error in assessments of portfolio credit risk, we start by calculating the five capital measures listed in Section 2.4 for 3000 large and as many small hypothetical portfolios. Even though they are constructed to have the same sectoral composition, the portfolios differ from each other with respect to the underlying risk parameters. For example, as reported in Table 3, the average PDs in large portfolios have a mean of 2.42% but can vary considerably across portfolios, ie from 1.79% to 3.12%. In comparison, the average asset return correlation changes little across portfolios, ranging between 9.14% and 10.73% for large portfolios. Interestingly, in line with a construct of the Basel II IRB approach, firms with higher PDs tend to be less correlated with the rest of the portfolios.[19] This is illustrated succinctly by the last line in each panel of Table 3, which reports negative correlations between individual PDs and the corresponding loading on a single common factor.[20]

Two of the capital measures deliver the target and shortcut capital levels, summary statistics of which are reported in Table 4. For large portfolios, the table reveals that target capital averages 2.95% (per unit of aggregate exposure) across the 3000 simulated portfolios. The corresponding shortcut level (based on a rule-of-thumb asset return correlation of 12%) is 76 basis points higher. The difference between the two capital levels amounts to 26% of target capital and reflects the fact that, compared to shortcut parameters, the parameters underpinning target capital imply 16-17 fewer defaults at a 99.9% credit VaR.[21] As we will

[19] The negative relationship between PDs and correlations (ie loading coefficients) is likely to be a general phenomenon. For example, Dev (2006) finds that global factors often play bigger roles for firms of better credit quality.

[20] This calculation is conducted under the one-factor approximation of the correlation matrix

[21] Because all exposures are equally weighted and have the same LGD (45%), one more default at the target VaR level raises the capital buffer for large (small) portfolios by 0.045 (0.225) percentage points, per unit of aggregate exposure.

see below, this result can change drastically if one changes the shortcut value of the asset return correlation.

Decomposing the discrepancy between target and shortcut capital for large portfolios reveals that errors caused by model misspecification play a minor role. Specifically, the multi-factor effect is with the expected negative sign (recall the discussion in Section 2.2.2) but entails a discrepancy that amounts to only 1% of the target capital level. This reveals that the one-factor approximation fits closely the raw correlation matrix.[22] Indeed, our one-factor approximation matches almost perfectly the level of average correlations (with a maximum discrepancy across simulated portfolios of less than 4 basis points) and explains on average 76% of the variability of pairwise correlations in the cross section of exposures.[23]

Similarly, the granularity effect is with the expected negative sign but, for large portfolios, leads to a small deviation from target capital. At about 4%, this deviation is only slightly higher than the one induced by the multi-factor effect. The magnitude of the granularity effect is largely in line with a granularity adjustment proposed by Gordy and Luetkebohmert (2006).[24]

By contrast, erroneous calibration of the ASRF model leads to much greater deviations from the target capital. For large portfolios, the correlation dispersion effect raises the capital measure by 35 basis points, which amounts to roughly 12% of the target level. The sign of the effect reflects the regularity that exposures with higher PDs tend to be less correlated with the rest of the portfolio. The shortcut capital measure ignores this regularity and, in line with the intuition provided in Section 2.3, overestimates the target capital.

The correlation level effect has a similarly important implication. Specifically, this effect reveals that raising the average correlation coefficient from 9.78% (the one observed in the data) to 12% leads to a 19% overestimation of the target capital level. The sign of the deviation is not surprising in light of the discussion in Section 2.3. It should be noted, however, that adopting another shortcut level for the average asset return correlation would change

[22] The discussion of the multi-factor effect in this paper should not be confused with the analysis in Duellmann and Masschelein (2006). These authors interpret the multi-factor effect as arising from a sectoral concentration in a bank's portfolio. Under such an interpretation, the multi-factor effect reflects not only the importance of multiple risk factors but also what we call here correlation level and dispersion effects.

[23] The goodness-of-fit measure for the one-factor approximation is described in Appendix B. Across the 3000 simulated large portfolios, this measure ranges between 67% and 85%. For small portfolios, this range is 63% to 86%.

[24] An application of the granularity adjustment formula (equation (6)) in Gordy and Luetkebohmert (2006) would match exactly a granularity effect that leads to a 5.4% underestimation of the target capital for large portfolios and a 24% underestimation for small portfolios.

the result dramatically. For instance, a shortcut level of 6% leads to a 32% *under*estimation of the target level.[25]

Turning to small portfolios, the decomposition results are qualitatively the same, with the notable exception of the granularity effect. In these portfolios, a much smaller portion of the idiosyncratic risk is diversified away and the granularity effect equals 53 basis points. This amounts to a 15% underestimation of the target level.

4.1.1 Regression analysis

The importance of correlation level and dispersion effects can be appreciated from a different point of view via regression analyses, the results of which are shown in Table 5. The regressions – run on the cross section of 3000 simulated portfolios – are simple linear models of *capital discrepancy*, which is defined as shortcut capital (based on a correlation of 12%) minus target capital. We include three blocks of explanatory variables. The first block consists of "pure" correlation characteristics: the average level and the dispersion in correlation coefficients for each simulated portfolio.[26] The second block of variables captures possible interaction between correlation coefficients and PDs and includes the product of average correlation and average PD and the correlation between PDs and loading coefficients, where the last variable is estimated under the one-common factor approximation. The third block includes our measure of how well the one-factor approximation explains the observed correlation matrix (defined in Appendix B). This measure reflects the importance of the multi-factor effect in explaining capital discrepancies. Unfortunately, it is impossible to include a variable that proxies for the granularity effect because, given that all 3000 simulated portfolios are homogeneous, there is no such variable that varies across portfolios.

In line with our previous findings, the regression results reveal that the correlation level and dispersion variables have strong explanatory power. These variables explain about 30% of the capital discrepancy and enter the regressions with statistically significant coefficients of the correct signs. The negative sign of the coefficient of the average correlation variable reflects the fact that a higher correlation in the data increases the target capital ratio and, thus, has a negative effect on capital discrepancy. In turn, the positive sign of the coefficient of the correlation dispersion variable is consistent with the above discussion that, when firms

[25] The rule-of-thumb asset return correlations reported in the literature range between 5 and 25%.

[26] This dispersion is calculated as the standard deviation of the loading coefficients under the one-factor approximation of the correlation matrix. Recall that the dispersion of shortcut correlations is zero by construction.

with high PDs are less correlated with the rest of the portfolio, lower correlation dispersion raises target capital. Since the shortcut capital abstracts from correlation dispersion, the correlation dispersion variable has a positive impact on capital discrepancy.

Moreover, the regressions also show that the interaction between heterogenous PDs and heterogenous correlations has a strong impact on assessments of portfolio credit risk. The two interactive variables are both statistically and economically significant, and including them in the regressions for large portfolios increases the adjusted R^2 by 44 percentage points to 76%. Similarly, for small portfolios, the adjusted R^2 increases by 34 percentage points to 59%.

Lastly, the goodness-of-fit of the one-factor model cannot explain any of the variability of capital discrepancy across simulated portfolios. The goodness-of-fit measure is statistically insignificant in the regression for small portfolios, and economically insignificant for either portfolio size. In addition, adding this variable to the regression does not affect adjusted R^2. This is just another illustration of our earlier finding that the multi-factor effect is negligible.

4.2 Estimation errors

The above results show that shortcut capital measures can deviate substantially from target. In practice, shortcut measures are likely to be adopted by less sophisticated users of the ASRF model who face constraints in terms of data and analytical skills. By contrast, large and complex financial institutions are likely to construct their own estimates of asset return correlations on the basis of in-house databases. This section demonstrates that, for realistic sizes of such databases, estimation errors in the correlation parameters are also likely to lead to large flaws in assessments of portfolio credit risk.

In order to quantify plausible estimation errors, we consider a portfolio whose "true" credit risk parameters match those of the "typical" portfolio in our data set. For this portfolio, we impose the simplifying assumption of homogeneous PDs (1%), LGDs (45%) and pairwise asset return correlations (9.78%) and consider different numbers of underlying exposures (see Table 6). Abstracting from issues of granularity, this assumption allows us to use the ASRF model and calculate that the desired capital buffer, dubbed "benchmark", equals 2.97% for each portfolio size. This is the typical (ie average) capital buffer across the simulated portfolios.

Then, we place ourselves in the shoes of a user of the model who does not know the exact asset return correlations but estimates them from available data. Specifically, we endow the user with 60, 120 or 300 months of asset returns data – drawn from the "true" underlying distribution – and calculate the sample correlation matrix. In order to quantify a plausible range of errors in the estimate of the correlation matrix, we repeat this exercise 1000 times. As reported in panels A and B of Table 6, the sample correlations contain estimation error that remains substantial even for 300 months (or 25 years) of data.

Panel C of the table reveals how estimation errors in correlation coefficients translate into deviation from the desired benchmark capital buffer.[27] First, these deviations are affected little by the number of exposures in the portfolio. Second, at standard confidence levels, the deviations can be large even for the longest time series considered. Taking a portfolio consisting of 1000 exposures and a user who has 120 months of data as an example, estimated capital buffers can deviate from the benchmark level by as much as 23% with a 95% probability. Finally, estimated capital buffers exhibit a positive bias relative to the benchmark level: their average level is invariably higher than 2.97%. This is because the true correlation structure is assumed to be homogenous, while small-sample errors introduce dispersion in estimated correlation coefficients. By the intuition presented in Section 2.3, this dispersion raises the implied capital buffer in the presence of homogenous PDs.

The bias in estimated capital vanishes if the user of the model knows the true correlation structure. In the present case, this means that the user knows that correlation coefficients are the same across all pairs of firms but still needs to estimate their exact value. Operationally, this translates into the user estimating the capital buffer on the basis of the average pairwise correlation in its sample. The results, based on the 1000 simulations (for each portfolio size) analysed above, are reported in panel D of Table 2.3. As expected, there is no longer bias in the estimated capital buffer but the plausible errors in this estimate are still substantial. Taking the same example as above, a user who has a portfolio of 1000 names and 120 months of data can be reasonably expected to estimate a capital buffer that is 20% lower or higher than the benchmark.

4.3 Alternative asset return distributions

It has been widely documented in the literature that asset returns are driven by distributions that have fatter tails than the convenient Gaussian distribution. This observation is not innocuous because, for example, a Gaussian assumption tends to bias capital buffers

[27] In order to focus on issues in the estimation of the interdependence of credit risk across exposures, we assume that the user knows the true PD and LGD.

downwards as long as the empirical distribution of asset returns is driven by fat tails in the distribution of the common factor. Such a distribution of the common factor implies a large probability that assets of several firms fall below the threshold values at the same time, ie a large probability of default clustering. We examine this issue in the remainder of the present subsection.

In order to focus on the impact of alternative asset return distributions on capital measures, we consider a homogeneous portfolio in which all PDs equal 1% and all correlation coefficients of asset returns equal 6%, 9.78% or 18%. For such a portfolio, we use the general ASRF formula (equation (3)) and derive capital buffers under different distributional assumptions. As reported in Table 7 (top half), we consider cases in which the single common factor of asset returns is distributed Student t (with various degrees of freedom) but the idiosyncratic factors are Gaussian. As expected, fatter tails in the distribution of the common factor (ie fewer degrees of freedom) translate into larger deviations from a capital buffer derived under the assumption that all variables are Gaussian. This deviation decreases when idiosyncratic factors are also allowed to follow a Student t distribution (bottom half of the table).

In order to decide what distributional assumptions are supported by the data, we consider time series of asset returns provided by Moody's KMV for all firms in our sample. Across the entire cross section of firms, the sample kurtosis (a measure of the fatness of the tails) has a mean of 7.28 and a median of 4.83.[28] Comparing these estimates to the kurtosis of asset returns implied by different distributional assumptions (Table 7, right panel) we see that there is evidence in the data for a "double-t" assumption, under which both common and idiosyncratic factors are distributed Student t with 5 or 7 degrees of freedom.

If the asset returns are indeed driven by such double-t distributions, then a Gaussian assumption, albeit convenient, would lead to substantial deviations from the desired capital buffer. For a specific example, consider the case in which the common asset return correlation is 9.78% (the average level in our sample). For such a portfolio, a double-t distribution (with 5 degrees of freedom for each factor) implies a capital buffer of 3.63% per unit of aggregate exposure, which is 22% higher than the capital based on a Gaussian assumption.

[28] These estimates are based on monthly data on asset values between August 2001 and July 2006.

5 Concluding remarks

This paper has quantified the relative importance of alternative sources of error in model-based assessments of portfolio credit risk. We have found that a misspecification of the popular ASRF model is likely to have limited impact on such assessments, especially for large well-diversified portfolios. By contrast, erroneous calibration of the model – driven by flaws in popular rule-of-thumb values of asset return correlation, plausible small-sample estimation errors, or a wrong assumption regarding the overall distribution of asset returns – can affect substantially measures of portfolio credit risk. These results highlight challenging tasks for credit risk managers and supervisory authorities, especially since our analysis has abstracted from several additional sources of error in risk parameter estimates. These sources relate to the estimation of individual PDs and LGDs, time variation in asset return correlations, and structural breaks in the credit environment that can impair the useful content of available data.

Appendix

A Monte Carlo simulations

Monte Carlo simulations deliver the target capital level. This method can be applied to any portfolio comprising N equally weighted exposures, provided that the associated probabilities of default, PD_i, losses-given-default, LGD, and correlation matrix of asset returns, R, are known.

The method consists of three general steps. In the first step, one uses the vector $\{PD_i\}_{i=1}^N$ and the assumption that asset returns are distributed as standard normal variables to obtain an $N \times 1$ vector of default thresholds. In the second step, one draws an $N \times 1$ vector from N standard normal variables whose correlation matrix is R. The number of entries in this vector that are smaller than the corresponding default threshold is the number of simulated defaults for the particular draw. In the third step, one repeats the second step a large number of times to derive the probability distribution of the number of defaults. Denoting this distribution's $(1-\alpha)^{th}$ percentile by β and the average PD in the portfolio by $\overline{PD}$, the target capital for a credit value-at-risk confidence level of $(1-\alpha)$ equals $LGD \cdot (\frac{\beta}{N} - \overline{PD})$ per unit of exposure.

In specific applications, we set $N = 200$ or 1000, $LGD = 45\%$, $\alpha = 0.1\%$ and $\{PD_i\}_{i=1}^N$ and R as estimated by Moody's KMV. In addition, the estimate of β relies on 500000 Monte Carlo simulations.

B Fitting a one-factor correlation structure

A one-factor approximation of an empirical correlation matrix is obtained as follows. Denote the empirical correlation matrix by Σ and its elements σ_{ij}, for $i, j \in \{1, \cdots, N\}$. The one-factor loading structure $\rho \equiv [\rho_1, ...\rho_N]$ that minimizes the discrepancies between the elements of Σ and their fitted counterparts are given by:

$$\min_{\rho} \sum_{i=1,\cdots,N-1} \sum_{j>i} (\sigma_{ij} - \rho_i \rho_j)^2$$

Andersen et al. (2003) propose an efficient algorithm to solve this minimization problem. The fitted correlation matrix $\hat{\Sigma}$ has elements $\rho_i \rho_j$.

We also construct a measure that reflects the "explanatory power" of the one-factor approximation:

$$\text{Goodness-of-fit measure} \equiv 1 - \frac{var(\epsilon)}{var(\sigma)}$$

where σ is a vector of all pairwise correlation coefficients σ_{ij} $(i, j = 1, \cdots, N, i < j)$ and ϵ is a vector of the errors $\sigma_{ij} - \rho_i\rho_j$ $(i, j = 1, \cdots, N, i < j)$. This measure reflects the degree to which the cross-sectional variation in pairwise correlations can be explained by common factor loadings in a single-factor framework.

C Gaussian copula

The Gaussian copula is an efficient algorithm for measuring portfolio credit risk when a portfolio consists of a finite number of exposures, the correlation matrix is driven by a factor-loading structure and underlying distributions are normal. The efficiency of the algorithm stems from the notion that, conditional on the realization of the common factor(s), defaults occurrences are independent across exposures. This allows for a closed-form solution for the conditional probability of joint defaults. The corresponding unconditional probability is then derived by integrating over the probability distribution of the common factor(s). For further detail, see Gibson (2004).

References

Andersen, Leif, Jakob Sidenius, and Susanta Basu (2003), "All your hedges in one basket," *Risk*, pages 67–72.

BCBS (2005), "An explanatory note on the Basel II IRB risk weight functions," *Basel Committee on Banking Supervision Technical Note*.

BCBS (2006), "Studies on credit risk concentration: an overview of the issues and a synopsis of the results from the Research Task Force project," *Basel Committee on Banking Supervision Research Task Force Working Paper*.

Berndt, Antje, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranz (2005), "Measuring default risk premia from default swap rates and EDFs," *Working Paper*.

Crosbie, Peter (2005), "Global correlation factor structure: modeling methodology," *Moody's KMV Documents*.

Crosbie, Peter and Jeffrey Bohn (2003), "Modeling default risk," *Moody's KMV White Paper*.

Das, Ashish and Shota Ishii (2001), "Methods for calculating asset correlations: a technical note," *KMV Documents*.

Dev, Ashish (2006), "The correlation debate," *Risk*, vol. 19.

Duellman, Klaus, Martin Scheicher, and Christian Schmieder (2006), "Asset correlations and credit portfolio risk – an empirical analysis," *Working Paper*.

Duellmann, Klaus (2006), "Measuring business sector concentration by an infection model," *Deutsche Bundesbank Discussion Paper*.

Duellmann, Klaus and Nancy Masschelein (2006), "Sector concentration in loan portfolios and economic capital," *Deutsche Bundesbank Discussion Paper*.

Emmer, Susanne and Dirk Tasche (2003), "Calculating credit risk capital charges with the one-factor model," *Journal of Risk*, vol. 7, 85–101.

Garcia Cespedes, Juan Carlos, Juan Antonio de Juan Herrero, Alex Kreinin, and Dan Rosen (2006), "A simple multifactor "factor adjustment" for the treatment of credit capital diversification," *Journal of Credit Risk*, vol. 2, 57–85.

Gibson, Michael (2004), "Understanding the risk of synthetic CDOs," *Finance and Economics Discussion Series*.

Gordy, Michael (2003), "A risk factor model foundation for ratings-based bank capital rules," *Journal of Financial Intermediation*, vol. 12, 199–232.

Gordy, Michael and Eva Luetkebohmert (2006), "Granularity adjustment for Basel II," *Working Paper*.

Heitfield, Erik, Steve Burton, and Souphala Chomsisengphet (2006), "Systematic and idiosyncratic risk in syndicated loan portfolios," *Journal of Credit Risk*, vol. 2, 3–31.

Hull, John and Alan White (2004), "Valuation of a CDO and an n-th to default CDS without Monta Carlo simulation," *Journal of Derivatives*, vol. 12, 8–23.

Loeffler, Gunter (2003), "The effects of estimation error on measures of portfolio credit risk," *Journal of Banking and Finance*, vol. 27, 1427–1453.

Longstaff, Francis, Sanjay Mithal, and Eric Neis (2005), "Corporate yield spreads: default risk or liquidity? New evidence from the credit-default-swap market," *Journal of Finance*, vol. 60, 2213–2253.

Lopez, Jose (2004), "The empirical relationship between average asset correlation, firm probability of default and asset size," *Journal of Financial Intermediation*, vol. 13, 265–283.

Martin, Richard and Tom Wilde (2002), "Unsystematic credit risk," *Risk*, pages 123–128.

Merton, Robert (1974), "On the pricing of corporate debt: the risk structure of interest rates," *Journal of Finance*, vol. 29, 449–470.

Morinaga, Satoshi and Yasushi Shiina (2005), "Underestimation of sector concentration risk by mis-assignment of borrowers," *Working Paper*.

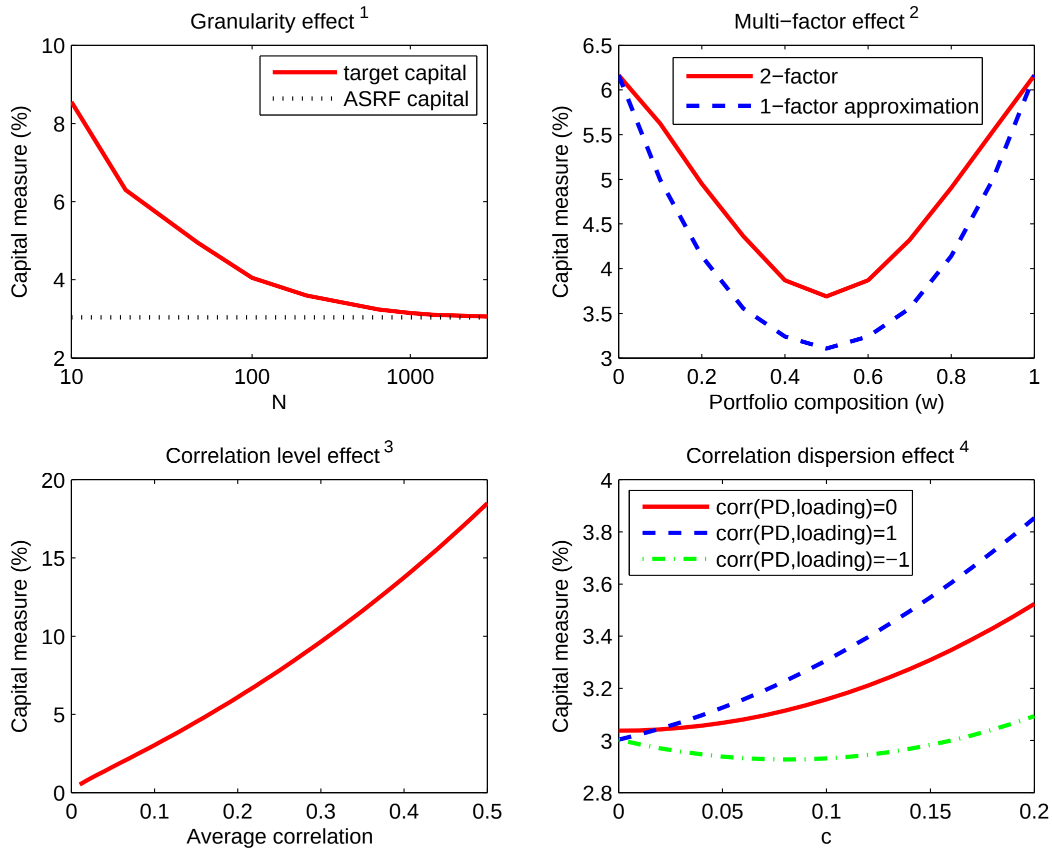
Pykhtin, Michael (2004), "Multi-factor adjustment," *Risk Magazine*, pages 85–90.

Tarashev, Nikola and Haibin Zhu (2006), "The pricing of portfolio credit risk," *Bank for International Settlements Working Paper*.

Vasicek, Oldrich (1991), "Limiting loan loss probability distribution," *KMV Working Paper*.

Vasicek, Oldrich (2002), "Loan portfolio value," *Risk*, pages 160–162.

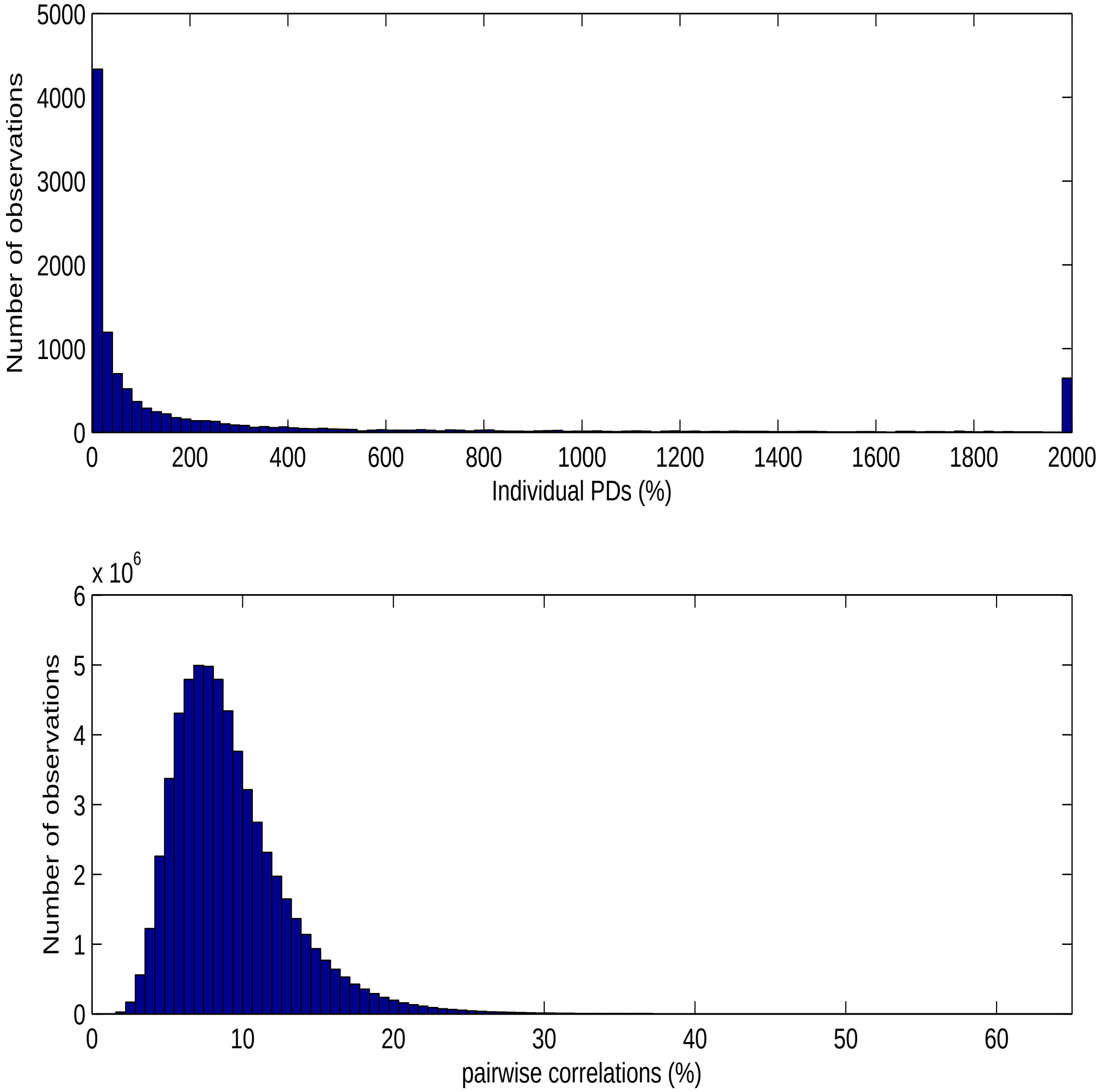
Figure 1: Four sources of error in capital measures



Note: Capital measures, in per cent and per unit of aggregate exposure, on the vertical axes. For each panel (unless noted otherwise), PD = 1% and LGD = 45% are the same across exposures.

[1] The solid line plots target capital for a portfolio in which all pairwise asset return correlations equal 10%. The number of exposures in the portfolio (N) varies across the horizontal axis. The dotted line plots the corresponding capital estimate when $N = \infty$. [2] The portfolio consists of two groups of exposures, with w denoting the weight of the first group. Within each group, the asset return correlation equals 20% for all exposure pairs. Inter-group correlation are zero. The solid line plots the target capital level, which incorporates the two common factors in the simulated data. The dashed line plots the capital calculated under a one-common-factor approximation of the correlation structure. This approximation imposes the same common-factor loading on all firms and does not affect the average asset return correlation. [3] Capital measures for different levels of the constant pairwise asset return correlation. [4] The solid line plots capital measures under the assumption that PD = 1%, there is a single common factor and the loadings on this factor are distributed uniformly in the cross section between $\sqrt{0.1} - c$ and $\sqrt{0.1} + c$. For the other two lines, PDs are distributed uniformly in the cross section between 0.5% and 1.5% and are perfectly positively (dashed line) or negatively (dot-dashed line) correlated with the common-factor loadings.

Figure 2: Distribution of individual PDs and pairwise correlations



Note: The parameter estimates are reported by Moody's KMV in July 2006 and relate to 10,891 non-financial firms.

Table 1: **Sectoral composition of simulated portfolios**

	Large portfolio		Small portfolio		*Memo:*
Sector	Number of names	Exposure weight (%)	Number of names	Exposure weight (%)	number of firms in the sample
Aerospace & Defense	31	3.1			105
Agriculture	9	0.9			56
Air Transportation	4	0.4			83
Apparel, Footwear, & Textiles	17	1.7			357
Automotive	51	5.1	19	9.5	198
Broadcast Media	43	4.3	16	8.0	191
Business Services	23	2.3			1132
Chemicals	42	4.2	15	7.5	940
Computer Equipment	10	1.0			746
Construction	43	4.3	16	8.0	277
Electric, Gas, & Sanitary	107	10.7	39	19.5	335
Electronics & Electrical	18	1.8			693
Entertainment & leisure	33	3.3			294
Fabricated Metals	17	1.7			146
Food, Beverages, & Tobacco	63	6.3	23	11.5	490
General Retail	31	3.1			133
Glass & Stone	6	0.6			149
Health care	38	3.8	14	7.0	178
Legal & Other Services	16	1.6			452
Lodging	22	2.2			70
Machinery & Equipment	36	3.6	13	6.5	645
Medical Equipment	10	1.0			334
Mining	6	0.6			486
Miscellaneous Manufacturing	18	1.8			130
Non-defense Trans. & Parts	2	0.2			53
Oil Refining & Delivery	24	2.4			100
Oil & Gas Exploration	55	5.5	20	10.0	458
Other Trans. Services	22	2.2			108
Paper & Forestry	23	2.3			172
Personal Services	7	0.7			31
Primary Metals	11	1.1			188
Printing & Publishing	28	2.8			186
Repair Services & Rental	13	1.3			37
Restaurants	9	0.9			141
Rubber & Plastics	18	1.8			120
Semiconductors	2	0.2			177
Telecommunications	69	6.9	25	12.5	177
Trucking & Warehousing	5	0.5			68
Water Transportation	5	0.5			104
Wood, Furniture, & Fixtures	13	1.3			151
Total	1000	100	200	100	10891
Name concentration index	0.0010		0.0050		
Sector concentration index	0.0432		0.1135		

Note: A concentration index is defined as the sum of squared weights, which are set either at the firm or the sector level.

Table 2: **Summary statistics (in percent)**

	mean	std. dev.	skewness	median	minimum	maximum
PDs	2.67	5.28	2.49	0.39	0.02	20.00
Pairwise correlations	9.24	3.86	1.87	8.45	0.29	65.00

Note: The sample includes 10,891 non-financial firms.

Table 3: **Characteristics of simulated loan portfolios (in percent)**

A. Large portfolios (1000 firms)					
	mean	std. dev.	median	minimum	maximum
average PD	2.42	0.19	2.42	1.79	3.12
std. dev. of individual PDs	5.16	0.26	5.16	4.25	6.14
median PD	0.26	0.03	0.26	0.18	0.36
average correlation	9.78	0.22	9.77	9.14	10.73
std. dev. of loadings	9.33	0.31	9.32	8.33	10.47
corr (PD, loadings)	-20.0	2.04	-20.1	-26.7	-12.8
B. Small portfolios (200 firms)					
	mean	std. dev.	median	minimum	maximum
average PD	2.28	0.36	2.26	1.24	3.68
std. dev. of individual PDs	5.05	0.53	5.06	3.01	6.89
median PD	0.24	0.05	0.23	0.11	0.55
average correlation	10.49	0.44	10.48	8.99	12.00
std. dev. of loadings	10.54	0.70	10.55	7.80	12.79
corr (PD, loadings)	-19.8	4.59	-20.2	-31.8	-1.2

Note: The results are based on 3,000 simulated portfolios and are carried out in two steps. First, portfolio-specific characteristics specified by row headings are calculated for each simulated portfolio. Second, summary statistics specified by column headings are calculated for each of the portfolio specific characteristics obtained in the first step. “Loadings” are estimated under a one-common-factor approximation of the correlation structure and refer to the firm-specific loadings of asset returns on the single common factor.

Table 4: **Capital measures and four sources of errors (in percent)**

A. Large portfolios (1000 firms)

	mean	std. dev.	median	95% interval	50% interval
Target capital [1]	2.95	0.16	2.95	[2.64, 3.27]	[2.84, 3.05]
Deviation from target due to: [2]					
Multi-factor effect [3]	-0.03	0.03	-0.045	[-0.09, 0]	[-0.045, 0]
Granularity effect [4]	-0.11	0.01	-0.11	[-0.14, -0.09]	[-0.12, -0.10]
Correlation dispersion effect [5]	0.35	0.04	0.35	[0.27, 0.43]	[0.32, 0.38]
Correlation level effect [6]	0.55	0.06	0.55	[0.44, 0.66]	[0.52, 0.59]
"shortcut" capital (corr=12%)	3.71	0.18	3.71	[3.37, 4.06]	[3.59, 3.83]
Memo: correlation level effect if:					
corr=6%	-0.96	0.07	-0.96	[-1.11, -0.83]	[-1.00, -0.91]
corr=18%	2.01	0.09	2.01	[1.84, 2.18]	[1.95, 2.07]
corr=24%	3.47	0.13	3.47	[3.23, 3.72]	[3.39, 3.56]

B. Small portfolios (200 firms)

	mean	std. dev.	median	95% interval	50% interval
Target capital [1]	3.35	0.30	3.34	[2.78, 3.94]	[3.15, 3.53]
Deviation from target due to: [2]					
Multi-factor effect [3]	-0.04	0.10	0	[-0.225, 0]	[0, 0]
Granularity effect [4]	-0.53	0.07	-0.53	[-0.65, -0.41]	[-0.59, -0.47]
Correlation dispersion effect [5]	0.38	0.11	0.37	[0.17, 0.58]	[0.30, 0.45]
Correlation level effect [6]	0.36	0.11	0.36	[0.15, 0.61]	[0.29, 0.44]
"shortcut" capital (corr=12%)	3.52	0.34	3.51	[2.85, 4.23]	[3.28, 3.75]
Memo: correlation level effect if:					
corr=6%	-1.07	0.12	-1.07	[-1.31, -0.85]	[-1.15, -0.99]
corr=18%	1.76	0.19	1.75	[1.41, 2.15]	[1.63, 1.88]
corr=24%	3.15	0.27	3.14	[2.65, 3.70]	[2.97, 3.33]

Note: Summary statistics for the simulated portfolios underpinning Table 3 (3,000 simulations for each portfolio size). The column entitled "95% interval " reports the 2.5th and 97.5th percentiles of the statistics specified in the particular row heading. The column entitled "50% interval" reports the corresponding 25th and 75th percentiles.

[1] Based on Moody's KMV estimates of PDs and asset return correlations and a Monte Carlo procedure for calculating the probability distribution of default losses. [2] Four sources of deviation from the target capital level; a negative sign implies underestimation. The sum of the target capital level and the four deviations equals the shortcut capital level. Each deviation (see table notes 3 to 6) is based on the assumptions underlying previous deviations plus one additional assumption. [3] For the multi-factor effect, the correlation matrix underpinning the target capital level is approximated under the assumption that there is a single common factor. [4] For the granularity effect, there is the additional assumption that the number of firms is infinite. [5] For the correlation dispersion effect, the additional assumption is that the loadings on the single common factor are the same across exposures. [6] For the correlation level effect, the additional assumption imposes a different, shortcut, level on the constant pairwise correlation.

Table 5: **Explaining the differences between target and shortcut capital**

	Large portfolios			Small portfolios		
	1	**2**	**3**	**1**	**2**	**3**
constant	0.025	0.024	0.024	0.019	0.017	0.017
	(54.2)	(86.2)	(86.1)	(34.4)	(40.7)	(28.7)
avg corr	-0.209	-0.278	-0.281	-0.182	-0.216	-0.216
	(7.5)	(65.9)	(63.7)	(26.0)	(41.4)	(41.0)
std dev of loading coefficients	0.037	0.039	0.035	0.020	0.024	0.023
	(51.9)	(13.4)	(10.3)	(4.5)	(7.2)	(4.9)
avg corr · avg PD		1.620	1.579		0.975	0.974
		(47.3)	(41.3)		(20.9)	(20.7)
corr(PD, loading coefficient)		-0.016	-0.016		-0.015	-0.015
		(51.8)	(51.2)		(39.6)	(39.5)
goodness-of-fit of one-factor model			0.001			0.0001
			(2.4)			(0.1)
adjusted R^2	0.32	0.76	0.76	0.25	0.59	0.59

Note: t-statistics in parentheses. The regression is based on 3,000 simulations for each portfolio size. The dependent variable equals shortcut capital (based on asset return correlation of 12%) minus target capital (see Table 4). Loading coefficients are estimated under a one-common-factor approximation of the correlation structure of asset returns. The goodness-of-fit of the one-common-factor approximation is measured as outlined in Appendix B.

Table 6: **Impact of estimation errors (in per cent)**

A. Sample average of pairwise correlations				
	N=100	N=200	N=500	N=1000
T=60	9.72 [6.5, 13.3]	9.67 [6.4, 13.3]	9.64 [6.6, 13.0]	9.63 [6.9, 12.7]
T=120	9.77 [7.4, 12.4]	9.77 [7.6, 12.1]	9.76 [7.6, 12.1]	9.72 [7.7, 12.0]
T=300	9.75 [8.3, 11.3]	9.79 [8.3, 11.3]	9.74 [8.4, 11.2]	9.77 [8.4, 11.25]

B. Sample standard deviation of loading coefficients				
	N=100	N=200	N=500	N=1000
T=60	12.11 [10.3, 14.1]	11.84 [10.5, 13.2]	11.65 [10.8, 12.5]	11.58 [10.8, 12.2]
T=120	8.50 [7.4, 9.8]	8.34 [7.5, 9.1]	8.17 [7.6, 8.8]	8.13 [7.7, 8.6]
T=300	5.35 [4.6, 6.2]	5.24 [4.7, 5.7]	5.16 [4.8, 5.5]	5.12 [4.9, 5.4]

C. Estimated capital, based on one-factor loading structure				
	N=100	N=200	N=500	N=1000
T=60	3.49 [2.6, 4.6]	3.44 [2.5, 4.5]	3.42 [2.6, 4.4]	3.41 [2.6, 4.3]
T=120	3.23 [2.6, 4.0]	3.22 [2.6, 3.9]	3.21 [2.6, 3.9]	3.19 [2.6, 3.9]
T=300	3.07 [2.7, 3.5]	3.07 [2.7, 3.5]	3.06 [2.7, 3.5]	3.07 [2.7, 3.5]
benchmark	2.97	2.97	2.97	2.97

D. Estimated capital, based on constant loading coefficients				
	N=100	N=200	N=500	N=1000
T=60	2.97 [2.1, 4.0]	2.95 [2.0, 4.0]	2.94 [2.1, 3.9]	2.94 [2.2, 3.8]
T=120	2.98 [2.3, 3.7]	2.98 [2.4, 3.7]	2.97 [2.4, 3.7]	2.96 [2.4, 3.6]
T=300	2.97 [2.6, 3.4]	2.98 [2.6, 3.4]	2.97 [2.6, 3.4]	2.97 [2.6, 3.4]
benchmark	2.97	2.97	2.97	2.97

Note: Results are based on 1000 simulations of the asset returns of N firms over T months. Each pair of asset return series is drawn from two standard normal variables with a correlation coefficient of 9.78%. Panel A reports means (across simulations) of the cross-sectional averages of pairwise sample correlations. In brackets, Panel A reports the corresponding 2.5th and the 97.5th percentiles of these averages. Panel B reports the same statistics for the cross-sectional standard deviations of sample pairwise correlations. Panels C and D report summary statistics of alternative capital measures, all of which are based on: (i) the ASRF formula (equation 4), (ii) the assumption that all exposures have the same PD of 1% and the same LGD of 45%, and (iii) alternative asset return correlations. The "benchmark" row reports the capital measure under the true correlation structure. The other three rows in Panel C report the mean (across simulations) of capital measures based on a one-common-factor approximation of the sample correlation matrices of simulated asset returns (these matrices underpin Panels A and B). In brackets, Panel C reports the 2.5th and the 97.5th percentiles of these capital measures. The corresponding rows in Panel D report corresponding statistics for capital measures underpinned by uniform pairwise correlations, equal to the average pairwise correlation in the simulated sample.

Table 7: **Alternative distributional assumptions**

	Capital measure (in per cent)			*Memo: kurtosis*		
	corr=6%	corr=9.78%	corr=18%	corr=6%	corr=9.78%	corr=18%
Student t						
(5,∞)	4.33	7.24	14.31	3.02	3.05	3.17
(7, ∞)	3.33	5.45	10.65	3.01	3.02	3.06
(10, ∞)	2.77	4.45	8.55	3.00	3.01	3.03
(15, ∞)	2.43	3.87	7.32	3.00	3.00	3.02
(20, ∞)	2.27	3.58	6.74	3.00	3.00	3.01
(5, 5)	2.00	3.63	9.08	8.01	7.65	6.97
(7, 7)	1.92	3.30	7.38	4.78	4.65	4.41
(10, 10)	1.91	3.16	6.59	3.90	3.83	3.71
(15, 15)	1.91	3.07	6.11	3.48	3.45	3.38
(20, 20)	1.91	3.04	5.92	3.33	3.31	3.26
Gaussian	1.92	2.97	5.45	3.00	3.00	3.00

Note: In the first column, the two numbers in parentheses refer to the degrees of freedom of the t-distribution of asset returns' common and idiosyncratic factor, respectively. The capital measures are obtained by applying the general ASRF formula (equation 3) to a homogeneous portfolio, in which PD=1% and LGD=45% for each exposure and the common pairwise asset return correlation coefficient is specified in the column heading. The default boundary for such exposures is calculated on the basis of 10 million simulations.