BANK FOR INTERNATIONAL SETTLEMENTS

BIS Working Papers

No 280

Measuring portfolio credit risk correctly: why parameter uncertainty matters

by Nikola A. Tarashev

Monetary and Economic Department

April 2009

JEL classification: G20, G32, C11.

Keywords: Correlated defaults; Estimation error; Risk management.

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

Copies of publications are available from:

Bank for International Settlements
Press & Communications
CH-4002 Basel, Switzerland

E-mail: publications@bis.org

Fax: +41 61 280 9100 and +41 61 280 8100

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)

ISBN 1682-7678 (online)

Measuring Portfolio Credit Risk Correctly: Why Parameter Uncertainty Matters*

Nikola A. Tarashev†

April 3, 2009

Abstract

Why should risk management systems account for parameter uncertainty? In order to answer this question, this paper lets an investor in a credit portfolio face non-diversifiable estimation-driven uncertainty about two parameters: probability of default and asset-return correlation. Bayesian inference reveals that – for realistic assumptions about the portfolio's credit quality and the data underlying parameter estimates – this uncertainty substantially increases the tail risk perceived by the investor. Since incorporating parameter uncertainty in a measure of tail risk is computationally demanding, the paper also derives and analyzes a closed-form approximation to such a measure.

JEL Classification: G20, G32, C11

Keywords: Correlated defaults; Estimation error; Risk management

*I would like to thank Rafael Schmidt for valuable help with some of the technical aspects of the paper. Michael Gordy, Kostas Tsatsaronis and seminar participants at the Bank for International Settlements also provided useful suggestions and comments on an earlier draft. This paper reflects my views only, not necessarily the views of the BIS.

†Bank for International Settlements, Centralbahnplatz 2, CH-4002 Basel, Switzerland. E-mail: nikola.tarashev@bis.org.

Introduction

Measures of credit risk are often based on an analytic model and on the assumption that the parameters of this model are known with certainty. In turn, it is common practice for risk management systems to rely on such measures because of their tractability, even though it is attained at the cost of ignoring **estimation noise**. This practice may impair severely the quality of risk management systems because, besides credit-risk factors, estimation noise is another important determinant of uncertainty about potential losses.

In order to substantiate this claim, this paper generalizes the popular asymptotic single risk factor (ASRF) model of portfolio credit risk by allowing for noisy estimates of two key parameters: probability of default (PD) and asset-return correlation.[1] Applied to a stylized empirical framework, the generalized model delivers two alternative measures of tail risk that help underscore the importance of estimation noise. The first measure is a **naive** value-at-risk (VaR) of the portfolio, which accounts for the credit-risk factor under the assumption that point estimates of the PD and asset-return correlation are equal to the true values of the respective parameters. The second is the **correct** VaR measure, which accounts not only for the credit-risk factor that influences the naive VaR but also for the uncertainty stemming from estimation noise.

A Bayesian inference procedure reveals that ignoring estimation noise leads to a substantial understatement of the correct VaR. In the benchmark specification (where an investor in a homogeneous portfolio estimates the PD and asset-return correlation at 1% and 20%, respectively, on the basis of data covering 200 obligors over 10 years) the correct VaR is 27% higher than the corresponding naive VaR. This result is striking, not least because the underlying stylized empirical framework incorporates a lower bound on the amount of estimation noise.

In addition, accounting for estimation noise (correctly) dampens the sensitivity of VaR measures to changes in parameter estimates. The flip side of this result is that, by abstracting from estimation noise, naive VaRs overstate the information content of parameter estimates. A concrete implication of the result is that the difference between the correct and naive VaRs – i.e. the **add-on** induced by parameter uncertainty – decreases (increases) by less than the naive VaR when changes in parameter estimates suggest lower (higher) tail risk of the portfolio.

In comparison to Gössl (2005) – which also argues for incorporating estimation noise in measures of portfolio tail risk – this paper conducts the analysis in a more transparent framework that allows for comparing in a straightforward fashion the importance of

[1] The paper abstracts from issues arising from **model** uncertainty. For an empirical analysis of the impact of model mis-specification on portfolio tail risk, see Tarashev and Zhu (2008a).

different sources of noise. Namely, a lengthening of the time series of the available data from 5 to 10 or from 10 to 20 years is seen to reduce the correct VaR add-on by a factor of two. In comparison, similar changes to the size of the cross section have a markedly smaller impact on the add-on. This finding is rooted in the standard assumption that credit risk is driven by asset returns that are serially uncorrelated but are correlated across obligors, which implies that increasing the time series of the data brings in more information than expanding the cross section.

The transparent framework of this paper also helps to analyze the trade-off between accuracy and reduction of the computational burden. This trade-off underscores the advantages of an approximate VaR measure, which exists in closed form, accounts for uncertainty about the PD and can be easily adjusted to reflect uncertainty about the correlation coefficient stemming from noise in observed asset returns. Given a judicious adjustment for such noise, this measure approximates the correct VaR quite well and alleviates substantially the computational burden.

That said, owing to the underlying Bayesian inference procedure, the computational burden is substantial for both the correct and the approximate VaR measures. This procedure is instrumental for capturing an important empirical regularity. Namely, over a realistic range of parameter values, higher levels of the PD and asset-return correlation are associated with greater noise in the associated estimates. This dependence between estimation noise and true parameter values, which is missed if one circumvents Bayesian inference, raises the probability that the true PD and correlation are bigger than their point estimates and, consequently, raises the correct VaR.

From a general point of view, this paper provides support to a call made by Borio and Tsatsaronis (2004) for including "measurement error information" in financial reporting. In deriving an ideal information set for attaining efficiency of the financial system, that paper emphasizes measurement error – a specific example of which is estimation error – as a piece of information that is of natural interest to risk managers and supervisors alike. From this perspective, the results derived below highlight specific scenarios in which knowledge of measurement error is indeed highly valuable as such error accounts for much of the uncertainty about potential credit losses.

The present paper differs in an important way from a number of recent articles – e.g. Löffler (2003), Tarashev and Zhu (2008a) and Heitfield (2008) – that have analyzed estimation noise in the context of portfolio credit risk.[2] These articles examine how errors

[2] Of these three papers, only Löffler (2003) analyzes uncertainty about PD on the basis of actual default data and does so via non-parametric bootstrap. See Lando and Skodeberg (2002) and Hanson and Schuermann (2006) for an extensive analysis of bootstrap approaches to the derivation of PD confidence intervals and Cantor et al (2008) for an application of such an approach to a large dataset. The analysis below, just like Heitfield (2008) and Tarashev and Zhu (2008a), circumvents the use of

in parameter estimates translate into errors in capital measures that ignore parameter uncertainty. However, the articles do not address the fact that parameter uncertainty should be a key input to appropriately constructed measures of portfolio credit risk. In terms of the terminology introduced in this paper, these articles quantify drawbacks of naive VaRs but do not derive correct VaRs.

The rest of the paper is organized as follows. Section 1 describes the model and then derives alternative measures of portfolio VaR. These measures are considered in the context of an empirical framework that is outlined in Section 2. In turn, Section 3 presents and analyzes the quantitative results. This section also derives a closed-form approximation to portfolio VaR that accounts rigorously only for uncertainty about the PD. Finally, Section 4 provides two extensions of the baseline analysis.

1 Stylized credit portfolio

The impact of parameter uncertainty on measures of tail risk is analyzed on the basis of a stylized credit portfolio. There are n exposures in this portfolio and all of them are of equal size, which is set to $1/n$. The analysis considers the limit, $n \to \infty$, in which the portfolio is referred to as asymptotic or "perfectly fine-grained".

In addition, all exposures exhibit ex ante homogeneous credit risk. This is captured by assuming that the value of the assets of each obligor i, $V_{i,\tau}$, reflects this obligor's credit condition and evolves in the following way:[3]

$$\ln(V_{i,\tau}) = \ln(V_{i,\tau-\Delta}) + \mu^*\Delta + \sigma^*\sqrt{\Delta}\left(\sqrt{\rho^*}M_\tau + \sqrt{1-\rho^*}Z_{i,\tau}\right) \qquad (1)$$

$$\text{where } M_\tau \sim N(0,1),\ Z_{i,\tau} \sim N(0,1),$$

$$Cov(M_\tau, Z_{i,\tau}) = 0,\ Cov(Z_{i,\tau}, Z_{j,\tau}) = 0 \text{ for all } i \text{ and } j \neq i,$$

and Δ denotes the period, in years, between two observations. The riskiness of obligor i is driven by a factor that is common to all obligors in the portfolio, M, and a factor that is specific to this obligor, Z_i. The factors M_τ and $\{Z_{i,\tau}\}_{i=1}^n$ are also serially uncorrelated. The drift of the asset value, the volatility of the asset value and the share of the common factor in this volatility are controlled by μ^*, $\sigma^* > 0$ and $\rho^* \in [0,1]$, respectively. These parameters are the same for all obligors in the portfolio.

Obligor i defaults if and only if $\ln(V_{i,\tau})$ is below some threshold, D_i^*. Default events are assumed to occur only at the end, $t \in \{1, 2, \cdots, T\}$, of non-overlapping

bootstrap methods by assuming that the functional form, albeit not the parameter values, of the data generating process is known.

[3] In this paper, "obligor" and "exposure" are used as close synonyms.

and adjacent one-year periods, which may be longer than the periods between two consecutive observations of the obligors' assets, i.e. $1 \geq \Delta$. Then, assuming that the loss-given-default on each exposure is unity, equation (1) implies that the loss on this portfolio over the next year t, $L_{n,t}$, equals:

$$
\begin{aligned}
L_{n,t} &= \sum_{i}^{n} U_{i,t}, \text{ where} \\
U_{i,t} &= \begin{cases} 1/n & \text{if } \sqrt{\rho^*} M_t + \sqrt{1-\rho^*} Z_{i,t} < \Phi^{-1}(PD^*) \\ 0 & \text{otherwise} \end{cases} \\
\Phi^{-1}(PD^*) &= \left(D_i^* - \ln(V_{i,t-1}) - \mu^*\right)/\sigma^*
\end{aligned}
\tag{2}
$$

and PD^* is the unconditional one-year probability of default, which is assumed to be the same across exposures (requiring that so is $D_i^* - \ln(V_{i,t-1})$). The expression $\Phi^{-1}(PD^*)$ is henceforth referred to as the (standardized) default point.[4]

An investor is interested in the maximum portfolio loss that can be incurred within a year with probability $(1-\alpha)$, i.e. in the one-year VaR at the $(1-\alpha)$ confidence level. It will be assumed that the investor knows how the portfolio loss is determined, i.e. knows the model in (2) and the distribution of the credit risk factors, M and Z_i. However, the investor has to estimate the homogeneous asset-return correlation, ρ^*, and probability of default, PD^*.

The remainder of this section outlines three alternative measures of portfolio VaR that make different uses of the information available to the investor. The first measure is the naive VaR, which treats the point estimates of the asset-return correlation and probability of default – denoted by $\hat{\rho}$ and $\widehat{PD}$ – as equal to the true parameter values. The second measure equals the correct VaR perceived by the investor. This measure also relies on the point estimates $\hat{\rho}$ and $\widehat{PD}$ but, in addition, incorporates the investor's uncertainty about the true parameter values. The third measure, considered only in passing, has the same functional form as the naive VaR but, instead of the estimates $\hat{\rho}$ and $\widehat{PD}$, incorporates conservative values of the asset-return correlation and the probability of default. These values, although related to the size of the investor's uncertainty about the true parameters, are determined in an ad-hoc way.

[4] In order to streamline the analysis, this setup abstracts from important aspects of portfolio credit risk. Unlike Gordy and Lütkebohmert (2007), for example, it does not address real-life departures from perfect granularity, which influence tail risk by increasing the role of exposure-specific factors. In addition, the setup rules out stochastic shocks to loss-given-default and exposure-at-default, which are modelled as correlated with the common default-risk factor by Kupiec (2008), and abstracts from cross-sectional dispersion of credit risk parameters, which is analyzed by Tarashev and Zhu (2008a).

1.1 Portfolio VaR under the ASRF model

In deriving the naive VaR measure, the uncertainty in parameter estimates is abstracted from, which reduces the above setup to a special case of the popular asymptotic single risk factor (ASRF) model. Then, results in Gordy (2003) imply that the naive VaR of the portfolio at the $(1-\alpha)$ confidence level equals the sum of exposure-specific expected losses, conditional on the common risk factor, M, being at the α^{th} quantile of its distribution.[5] Given that the forecast horizon is one year, expression (2) implies that this boils down to:

$$\begin{aligned} VaR^{naive} &= E\left(L_n | M = \Phi^{-1}(\alpha), PD^* = \widehat{PD}, \rho^* = \hat{\rho}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(\widehat{PD}) - \sqrt{\hat{\rho}}\Phi^{-1}(\alpha)}{\sqrt{1-\hat{\rho}}}\right) \end{aligned} \tag{3}$$

where $\Phi(\cdot)$ denotes the standard normal CDF. In addition, time subscripts have been suppressed in order to alleviate the notation.

1.2 Portfolio VaR with parameter uncertainty

Since it abstracts from estimation noise, the naive VaR measure in (3) fails to reflect the fact that the correct VaR, i.e. the one perceived by the investor, incorporates candidate values of the PD and asset-return correlation, PD^c and ρ^c, as random variables. The joint probability density of these random variables is assumed to be well-defined, continuous and bounded away from zero everywhere on its support, which is $(0,1) \times (0,1)$. The next section presents a concrete empirical framework, in which this assumption is borne out.

Parameter uncertainty implies that the investor faces multiple common risk factors. These are the common credit-risk factor, M, and the common "estimation-risk" factors, PD^c and ρ^c. Since M is serially independent and the uncertainty embedded in ρ^c and PD^c is driven by past data, M is independent of ρ^c and PD^c.

Importantly, the presence of multiple risk factors violates a key assumption of the ASRF model, implying that the formula in (3) can no longer be used for calculating the VaR of the portfolio. Instead, it is necessary go through the following four steps. First,

[5] In general terms, the q^{th} quantile of the distribution of a generic random variable Y is defined as $Q_q(Y) \equiv \inf\{y : \Pr(Y \leq y) \geq q\}$.

consider the following conditional expected loss on the portfolio:

$$E\left(L_n|M, PD^c, \rho^c\right) = \sum_{i=1}^{n} \Phi\left(\frac{\Phi^{-1}\left(PD^c\right) - \sqrt{\rho^c}M}{\sqrt{1-\rho^c}}\right) \equiv E\left(L|M, PD^c, \rho^c\right) \tag{4}$$

which is a random variable that is independent of the sample size, n, owing to the assumed homogeneity of parameters across obligors. Second, denote the $(1-\alpha)$ quantile of this conditional expectation by $Q_{1-\alpha}\left(E\left(L|M, PD^c, \rho^c\right)\right)$. Third, denote the correct VaR of the portfolio at the $(1-\alpha)$ confidence level by $Q_{1-\alpha}\left(L_n\right)$. Finally, the following proposition, which is proved in Appendix A, states that the correct VaR of an asymptotic portfolio equals $Q_{1-\alpha}\left(E\left(L|M, PD^c, \rho^c\right)\right)$:

Proposition 1 As $n \to \infty$, $Q_{1-\alpha}\left(L_n\right) \to Q_{1-\alpha}\left(E\left(L|M, PD^c, \rho^c\right)\right)$.

It is important to note that there would generally not be an analytic expression for the correct VaR measure, even in the limit $n \to \infty$. This might make it tempting to consider alternative measures that focus directly on conservative values of the credit- and estimation-risk factors. Specifically, given the confidence level of the desired VaR, such an alternative could condition on the α quantile of the common credit factor and the $(1-a)$ quantiles of the probability of default and correlation candidates (henceforth, $PD^c_{1-\alpha}$ and $\rho^c_{1-\alpha}$):

$$\begin{aligned} VaR^{alt} &= E\left(L|M = \Phi^{-1}(\alpha), PD^c = PD^c_{1-\alpha}, \rho^c = \rho^c_{1-\alpha}\right) \\ &= \Phi\left(\frac{\Phi^{-1}\left(PD^c_{1-\alpha}\right) - \sqrt{\rho^c_{1-\alpha}}\Phi^{-1}(\alpha)}{\sqrt{1-\rho^c_{1-\alpha}}}\right) \end{aligned} \tag{5}$$

Albeit computationally more efficient, VaR^{alt} turns out to be larger than the correct VaR, $Q_{1-\alpha}\left(E\left(L|M, PD^c, \rho^c\right)\right)$. There are two reasons for this. First, in all the numerical examples considered below, the two alternative measures coincide when the credit-risk factor, M, is perfectly correlated with the estimation-risk factors, PD^c and ρ^c.[6] Second, since these risk factors are not perfectly correlated, there are "diversification benefits" that depress the correct VaR measure but do not affect VaR^{alt}. The difference between the two alternative measures is quantified in Sections 3.1 and 3.2 below.

[6] In this context, "perfectly correlated" is to be understood as follows: conditional on the common credit-risk factor being equal to the q^{th} quantile of its distribution, the PD and correlation candidates are at the $(1-q)^{th}$ quantiles of their respective distributions with probability 1.

2 The Empirical framework

This section outlines the derivation of the correct VaR measure, $Q_{1-\alpha}\left(E\left(L|M, PD^c, \rho^c\right)\right)$. In addition to the distribution of the common credit-risk factor, this measure incorporates the joint probability distribution of estimation-risk factors, which reflects features of the data used for estimation and the inference procedure. These features are assumed to be such as to ensure that the correct VaR measure entails reasonable computational burden and, at the same time, is informative about the impact of estimation noise in real-life applications.

2.1 Stylized dataset

The investor observes asset returns and default rates, which are delivered by the data-generating process specified in (1) and (2). It will be assumed that these data cover T cohorts, each one of which comprises N obligors. Each cohort is followed for one year $t \in \{1, \cdots, T\}$ and, for each month in this year, the investor observes the assets of the N obligors (implying that the length of the period between two consecutive asset observations is $\Delta = 1/12$). At the end of each year t, the investor also observes the default rate in the respective cohort. In line with the investor's portfolio, all obligors in each cohort are characterized by an asset-return correlation that equals ρ^* and (at the beginning of the relevant year) a probability of default that equals PD^*.

The stylized dataset warrants four remarks. First, the assumed frequencies of asset and default rate observations are intended to capture common practice. Reportedly, in order to filter out high-frequency noise, practitioners base their estimates of asset-return correlations on weekly or monthly time series of assets that are obtained from daily equity prices and/or CDS spreads.[7] In turn, yearly observations of one-year default rates are standard ex post measures of short-term credit risk.

Second, despite the maintained focus on an asymptotic portfolio, the numerical results below reflect datasets with finite cross sections. This corresponds to the likely real-life case in which the investor has access to data on only a subset of his/her exposures. That said, the benchmark numerical exercise in this paper uses a portfolio of $N = 200$ exposures and it turns out that expanding such a cross section leads to small declines of estimation noise that have a limited impact on the VaR perceived by the investor (see Sections 3.1 and 3.2 below).

Third, the benchmark exercise incorporates one-year default rates realized over $T = 10$ years and asset returns realized over $T = 120$ months. To put the length of these

[7] In line with such practice, Moody's KMV publishes monthly estimates of the market value of the assets of the firms in its database. See Heitfield (2008) and Tarashev and Zhu (2008b) for further detail.

time series into perspective, note that the Basel II accord requires regulated institutions to base their PD estimates for corporate exposures on at least 5 years of data.

Fourth, changes in credit-risk outlooks are a frequent real-life phenomenon, which might invalidate the assumption that a fixed number of obligors, N, feature the same parameters, ρ^* and PD^*, over several years. A time invariant N is, however, in line with the illustrative nature of the analysis.

2.2 Bayesian inference procedure

The investor tackles estimation noise on the basis of a Bayesian inference procedure. Namely, the investor does not know the true values of the asset-return correlation and the probability of default, ρ^* and PD^*, but holds prior beliefs about the probability distribution of the candidate values, ρ^c and PD^c. Then, on the basis of the data described in Section 2.1 and knowledge of the true structure of the data generating process, the investor derives the point estimates, $\widehat{\rho}$ and $\widehat{PD}$, and uses them in order to update the prior beliefs into posterior probability distributions of ρ^c and PD^c.

More concretely, the inference procedure is conducted in two consecutive steps. This procedure reflects the fact that, in the adopted empirical framework, default-rate data do not affect the inference about the asset-return correlation but inference about the probability of default needs to condition on information about this correlation. The first step in the procedure derives the posterior distribution of correlation candidates, ρ^c, as follows:

- The investor starts with the prior belief that the PDF of ρ^c is $g(\rho^c)$.
- The investor then obtains the point estimate $\widehat{\rho}$ solely on the basis of data on asset returns. Given that asset returns are observed directly,[8] default data do not provide any additional information regarding the asset-return correlation.
- Recognizing that $\widehat{\rho}$ is affected by estimation noise, the investor uses the conditional PDF of the correlation estimate, $f(\widehat{\rho}|\rho^c)$,[9] in order to derive the posterior distribution of correlation candidates:

$$h(\rho^c|\widehat{\rho}) = \frac{g(\rho^c)\, f(\widehat{\rho}|\rho^c)}{\int g(\rho^c)\, f(\widehat{\rho}|\rho^c)\, d\rho^c} \tag{6}$$

[8] This assumption is relaxed in Section 4.2 below.

[9] As shown in Appendix B, this conditional PDF and its analog in the context of PD estimation are unaffected by the values of the drift and volatility parameters, μ^* and σ^* (recall expression (1)), but are affected by the fact that σ^* needs to be estimated.

The second step of the inference procedure derives the posterior distribution of the probability-of-default candidates, PD^c:

- The investor starts with the prior belief that the PDF of PD^c is $\tilde{g}(PD^c)$.
- Conditioning on a particular correlation candidate, ρ^c, the investor obtains the point estimate, $\widehat{PD}(\rho^c)$, solely on the basis of data on default rates.
- Then, using the conditional PDF of the probability-of-default estimate, $\tilde{f}\left(\widehat{PD}(\rho^c)|PD^c,\rho^c\right)$, the investor derives the conditional posterior distribution of candidate values as follows:

$$\tilde{h}\left(PD^c|\widehat{PD}(\rho^c),\rho^c\right) = \frac{\tilde{g}(PD^c)\tilde{f}\left(\widehat{PD}(\rho^c)|PD^c,\rho^c\right)}{\int \tilde{g}(PD^c)\tilde{f}\left(\widehat{PD}(\rho^c)|PD^c,\rho^c\right)dPD^c} \tag{7}$$

The investor is now in a position to derive the distribution of the conditional expected loss on the portfolio, $E(L|M,PD^c,\rho^c)$. He/she does so by recognizing that the expressions in (6) and (7) define the joint posterior distribution of ρ^c and PD^c and recalling that the credit risk factor, M, is a standard normal variable that is independent of ρ^c and PD^c.

2.3 Posterior distributions

In practical applications, a VaR measure would typically be based on a specific parameterization of the posterior distributions in (6) and (7). This paper adopts the following parameterization and assumes that it is known to the investor:

- The estimates of the asset-return correlation and default probability are delivered by minimum-variance unbiased estimators. Thus, given candidate parameter values, ρ^c and PD^c, the standard deviations of the noise in the point estimates, $\widehat{\rho}$ and $\widehat{PD}(\rho^c)$, equal the respective Cramer-Rao lower bounds, denoted by $\sigma_\rho(\rho^c)$ and $\sigma_{PD}(PD^c,\rho^c)$.[10]
- The conditional distributions of the estimators, $f(\bullet|\rho^c)$ and $\tilde{f}(\bullet|PD^c,\rho^c)$, are specified as follows:

[10] See Appendix B for further information on $\sigma_\rho(\rho^c)$ and $\sigma_{PD}(PD^c,\rho^c)$.

– $f(\bullet|\rho^c)$ is a beta PDF with a mean ρ^c and variance $\sigma^2_\rho(\rho^c)$:[11]

$$\begin{aligned} f(\bullet|\rho^c) &= beta(\bullet; A, B), \text{ where} \\ A &\equiv \frac{\rho^c}{1-\rho^c}B,\ B \equiv \frac{(1-\rho^c)\left(\rho^c(1-\rho^c)-\sigma^2_\rho(\rho^c)\right)}{\sigma^2_\rho(\rho^c)} \end{aligned}$$

– $\tilde{f}(\bullet|PD^c, \rho^c)$ has a mean PD^c and variance $\sigma^2_{PD}(PD^c, \rho^c)$. The exact shape of this distribution is such as to guarantee that the associated posterior of PD^c is as defined below.

- The posteriors of the parameter candidates, $h(\bullet|\widehat{\rho})$ and $\tilde{h}\left(\bullet|\widehat{PD}(\rho^c), \rho^c\right)$, are defined as follows:
  - $h(\rho^c|\widehat{\rho}) = f(\widehat{\rho}|\rho^c)$, as implied by (6) under a uniform (or diffuse) prior, $g(\rho^c) = 1$ for $\rho^c \in [0, 1]$.
  - $\tilde{h}\left(\bullet|\widehat{PD}(\rho^c), \rho^c\right)$ satisfies two criteria. First, in accordance with (7), its mean and variance are equal to the ones implied by $\tilde{f}(\bullet|PD^c, \rho^c)$ and a uniform prior, $\tilde{g}(PD^c) = 1$ for $PD^c \in [0, 1]$. Second, the implied posterior distribution of the default point $\Phi^{-1}(PD^c)$ is normal.

2.4 Measuring portfolio VaR

In quantifying portfolio VaR, the investor makes use of the following three pieces of information:

- a panel dataset of asset-returns and default rates;
- the posterior distribution of correlation candidates, $h(\bullet|\bullet)$;
- the conditional posterior distribution of PD candidates, $\tilde{h}(\bullet|\bullet, \bullet)$.

Then, the probability distribution of the conditional expected loss, $E(L|M, PD^c, \rho^c)$, is quantified via the following simulation procedure:

1. The data on asset returns produce a point estimate of their correlation, $\widehat{\rho}$.
2. A candidate ρ^c is drawn from the posterior distribution $h(\cdot|\widehat{\rho})$.[12]

[11] Note that the Fischer information inequality, which is behind the Cramer-Rao lower bounds on estimation noise, specifies only the first two moments of minimum-variance unbiased estimators.

[12] For the numerical exercise below, this distribution is discretized as described in Appendix C.

3. Conditioning on ρ^c, the point estimate of the probability of default, $\widehat{PD}(\rho^c)$, is obtained from data on default rates via maximum likelihood estimation.[13]

4. Continuing to condition on ρ^c, a value of $\Phi^{-1}(PD^c) - \sqrt{\rho^c}M$ is drawn from the normal PDF implied by the posterior distribution of PD^c, $\tilde{h}\left(\cdot|\widehat{PD}(\rho^c), \rho^c\right)$, the standard normal distribution of M and the independence between PD^c and M.

5. Repeating step 4 a large number of times (the specific exercise employs 1 million such repetitions) delivers the distribution of the conditional expectation $E(L|M, PD^c, \rho^c)$ for the given ρ^c.

6. Repeating steps 2–5 delivers the distribution of this conditional expectation for any point estimate $\widehat{\rho}$ and observed default rates.

By Proposition 1, the $(1-\alpha)$ quantile of the so-derived distribution equals the correct VaR at the $(1-\alpha)$ confidence level.

2.5 Discussion of the empirical framework

This sub-section revisits important aspects of the stylized empirical framework. First, some of the framework's underlying assumptions – e.g. homogeneity of exposures, a time-invariant size of the dataset, a convenient parameterization of probability distributions – are key for deriving measures of portfolio VaR that are both influenced by parameter uncertainty and not prohibitively burdensome to calculate. The reason is that these assumptions limit the number of the parameters of interest to two – i.e. the common PD and asset-return correlation – and help circumvent inference about "nuisance" parameters, which are not central to the problem at hand (such as the drift and volatility of asset returns).[14] To see the importance of this implication, note first that the calculation of a single VaR measure under the adopted framework requires roughly six days of computer time (on a Pentium(R) 4 CPU 3.20GHz machine with 2GB of RAM). Moreover, as indicated by the procedure described in Section 2.4, the number of simulations grows exponentially in the number of parameters relevant for deriving a VaR.

A second important aspect of the framework is that some of the simplifying assumptions are likely to depress the perceived VaR. For example, the assumptions of

[13] Most of the numerical analysis below makes use of the Cramer-Rao lower bounds on the variance of parameter estimators, keeping in the background the exact specification of these estimators. That said, simulation exercises reveal that the maximum-likelihood estimator of the probability of default attains the Cramer-Rao lower bound and is largely consistent with the assumed shape of $\tilde{f}(\bullet|PD^c, \rho^c)$. This estimator is referred to explicitly only in Section 3.3, which discusses how ρ^c affects $\widehat{PD}$.

[14] See Heitfield (2008) for a setup, in which it is necessary to make inference about nuisance parameters.

minimum-variance unbiased estimators, homogeneous exposures and a fixed size of the cross section of the data limit significantly the amount of estimation noise that is allowed to affect the investor's perception of risk. As a result, the correct VaR derived under such simplifying assumptions should be treated as a lower bound on the VaRs faced by investors in real-life portfolios, where many of these assumptions are violated.

That said, a third important aspect of the framework relates to the inference procedure, which has been designed to insulate the precision of parameter estimates from some highly stylized aspects of the dataset. Namely, when estimating the probability of default, the investor is not allowed to use information that is contained in observed asset returns but is missed by the estimate of their correlation (see Section 2.2). Thus, the investor is not allowed to exploit the fact that, since all obligors are assumed to be ex ante homogeneous, the minimum asset value among surviving obligors and the maximum asset value among defaulting obligors bound from below and above the possible values of the default point. Given the sizes of the datasets examined in this paper, the two bounds are typically so close to each other as to effectively **reveal** the default point and, thus, the PD.

The inference procedure adopted here abstracts from this clearly unrealistic implication. Indeed, Heitfield (2008) reports that, in real-life applications – where (i) obligors are heterogeneous and (ii) asset-return and default-rate data cover different sets of obligors – the common practice is to use asset-returns data to estimate the **average** correlation and to use the default data to estimate the **average** PD but **not** to use asset-returns data for a direct estimation of the default point. The empirical framework in this paper emulates this common practice.

A fourth important aspect of the empirical framework relates to the assumption that prior beliefs about the credit-risk parameters are diffuse. Although this assumption is quite in line with the level of generality in this paper, it could easily be replaced with other similarly acceptable alternatives. Section 4.1 below derives such alternatives on the basis of long historical data on default rates and then studies their implications for portfolio VaR.

Fifth, the assumption that asset returns are observed directly masks an important challenge associated with correlation estimates. In practice, asset returns are subject to observation noise because they need to be extracted from other variables, such as stock prices and CDS spreads. As shown in Section 4.2, abstracting from observation noise in asset returns can bias the VaR measure. The section also discusses conditions under which the bias can be positive or negative, and proposes a correction.

3 Results

The main result is that ignoring parameter uncertainty leads to a significant understatement of the portfolio tail risk perceived by the investor. Importantly, this result is derived within a stylized empirical framework, which, as argued above, provides only the *lower* bound of the impact of parameter uncertainty on VaR measures. In addition, the result is robust to changes in the point estimates of the credit risk parameters, $\widehat{PD}$ and $\widehat{\rho}$, and in the size of the dataset, N and T.

A numerical example, drawn from the top right-hand panel of Table 1, helps fix ideas. Consider the benchmark case in which an investor obtains point estimates $\widehat{PD} = 1\%$ and $\widehat{\rho} = 20\%$ on the basis of panel data on (monthly) asset-returns and (yearly) default rates that are observed for $T = 10$ years and $N = 200$ obligors. If the investor is interested in the portfolio VaR at the 99.9% confidence level and ignores estimation noise, he/she uses equation (3) and calculates a *naive* VaR that equals 14.55 cents on the dollar. However, applying Proposition 1 reveals that estimation noise requires an "add-on" – on top of the naive VaR – that equals 3.93 cents on the dollar. Thus, the *correct* VaR perceived by the investor is 27% higher than the naive one.

A comparison between the top and bottom right-hand panels of Table 1 reveals that the correct VaR measure is less sensitive to changes in the parameter estimates than the naive VaR. Changing the PD estimate in the benchmark example to $\widehat{PD} = 5\%$ results in the correct VaR measure rising from 18.48 (i.e. $14.55 + 3.93$) cents on the dollar to 43.37 ($38.44 + 4.93$) cents, or by a factor of 1.35. At the same time the naive VaR rises by a factor of 1.64, from 14.55 to 38.44 cents on the dollar. Further, this difference in sensitivities to changing parameter estimates becomes more pronounced as the sample size (N and/or T) decreases and, thus, parameter uncertainty increases.

From a different perspective, either reducing the cross section of the data or shortening its time series leads to higher uncertainty-induced VaR add-ons but the effect of the latter is considerably more important (see Table 1). Starting with the benchmark example and then halving the length of the time series (to $T = 5$ years of data) more than doubles the VaR add-on (to 8.20 cents on the dollar). By contrast, decreasing the size of the cross section by a factor of four (to $N = 50$ obligors) raises the VaR add-on by 50% (to 5.85 cents).

The remainder of this section analyzes in some detail the impact of parameter uncertainty on the VaR measure and, in the process, provides explanations for the above results. The analysis first considers uncertainty stemming only from the estimation of the asset-return correlation, under the assumption that the PD is observed directly. The advantage of focusing on this case stems from the existence of an analytical expression for the dependence of correlation uncertainty on the size of the dataset and the true parameter values. Then, the section proceeds to focus on the impact of uncertainty about the PD, assuming that the asset-return correlation is observed directly. Even though PD uncertainty cannot be analyzed analytically, there is a closed-form expression for the VaR that incorporates a measure of this uncertainty. The last part of this section argues in favour of using this closed form expression in order to approximate the correct VaR perceived by the investor.

3.1 Correlation uncertainty

Uncertainty about the asset-return correlation leads to a small VaR add-on. To parallel the above benchmark example, suppose that the investor knows with certainty that the true $PD^* = 1\%$ and obtains the point estimate $\widehat{\rho} = 20\%$ on the basis of $T = 120$ months of data covering the asset returns of $N = 200$ obligors. As reported in the top middle panel of Table 1, the resulting VaR add-on is 0.41 cents on the dollar. This is roughly 2.8% of the naive VaR, which ignores parameter uncertainty altogether, and slightly more than 10% of the add-on that incorporates noise in both the correlation and PD estimates.

Further numerical results, reported in Table 2, allow for analyzing the importance of alternative drivers of correlation uncertainty. For example, it can be seen that the VaR add-on induced by noise in the correlation estimate depends little on the size of the cross-section, N, but is quite sensitive to the length of the sample period, T. As reported in the centre panel of Table 2, shrinking the cross-section by a factor of 4 (i.e. switching from $N = 200$ to $N = 50$) leaves the correct add-on virtually unchanged when $T = 120$ months and raises it by 9% if $T = 60$.[15] By contrast, shortening the time series by a factor of 2 (i.e. switching from $T = 120$ to $T = 60$) roughly doubles the VaR add-on for both $N = 50$ and $N = 200$.

Heuristically, this result reflects the fact that, as the sample size expands, the convergence of estimates to the true parameter values is "slowed down" by correlation among observations. This convergence is underpinned by the Law of Large Numbers (LLN), which, in the present context, "works" fully in the time dimension (as there is serial independence) but only up to a point in the cross section (because the common credit-risk factor leads to a positive correlation among obligors). The intuition behind LLN is that, when the sample size increases, there is greater chance that estimation noise is "averaged out" and, as a result, point estimates become more precise. Of course,

[15] A "correct" add-on in Table 2 is one that incorporates fully correlation uncertainty but also reflects the maintained assumption that the PD is known. Similar definitions apply to the correct add-ons in Tables 3 and 5.

increasing either N or T leads to an averaging out of obligor-specific noise. By contrast, the noise stemming from the common factor is averaged out by increasing T but **not** by increasing N.

This intuition finds its concrete expression in the Cramer-Rao lower bound on the variance of the noise in correlation estimates (see Appendix B):

$$\sigma_{\rho}^{2}(\rho^{*};N,T)=\frac{2(1-\rho^{*})^{2}(1+(N-1)\rho^{*})^{2}}{TN(N-1)} \tag{8}$$

This variance decreases to zero as $T\rightarrow\infty$ but to a positive number, i.e. $2(1-\rho^{*})^{2}(\rho^{*})^{2}\big/T$, as $N\rightarrow\infty$. This is illustrated in Figures 1 and 2, which plot, respectively, the impact of increasing the time series and the cross-section of the data on $\sigma_{\rho}(\rho^{*};N,T)$.

A comparison among the top, middle and bottom panels of Table 2 reveals that the correct VaR add-on (driven only by noise in the correlation estimate) increases as the point estimate of correlation increases within the considered range. If $N=200$ and $T=120$ months, for example, this add-on increases from 0.26 to 0.41 and then to 0.50 cents on the dollar as the point estimate, $\widehat{\rho}$, increases from 10% to 20% and 30%. The reason for this is that the variance of the noise in correlation estimates increases as the true correlation increases from 0% to (roughly) 50% (see Figures 1 and 2). More precisely, by equation (8),[16]

$$\frac{d\sigma_{\rho}^{2}(\rho^{*};N,T)}{d\rho^{*}}>0 \text{ for } \rho^{*}\in\left[0,0.5\frac{N-2}{N-1}\right] \tag{9}$$

This points to a pitfall in measuring VaR without taking into account the dependence of the size of estimation uncertainty on the true value of the correlation. If this dependence were ignored, the posterior distribution of correlation candidates would coincide with $f(\bullet|\rho^{*}=\widehat{\rho})$, which is the PDF of the correlation estimator when the true parameter, ρ^{*}, happens to be at the point estimate, $\widehat{\rho}$. In this case, the Bayesian updating procedure would be redundant and, thus, the computational burden would be substantially reduced. However, as indicated by expression (9), a higher ρ^{*} is associated with greater dispersion in $f(\bullet|\rho^{*})$. Thus, a given point estimate, $\widehat{\rho}$, is more likely to be associated with $\rho^{*}>\widehat{\rho}$ than with $\rho^{*}<\widehat{\rho}$, implying that the posterior distribution of correlation candidates, $h(\bullet|\widehat{\rho})$, has a more pronounced right skew and a higher mean than $f(\bullet|\rho^{*}=\widehat{\rho})$. This result holds irrespective of the fact that the estimator of the asset-return correlation is unbiased.

Thus, ignoring that estimation noise depends on the true parameter value leads to

[16] This result holds true even if pairwise asset-return correlations are heterogeneous.

a mis-specified posterior distribution that attributes too much (little) probability mass to low (high) correlation candidates. In turn, since higher values of the asset-return correlation imply higher tail risk, this mis-specification leads to an understatement of the VaR perceived by the investor. The magnitude of such an understatement is illustrated in Table 2 by the difference between correct VaR add-ons and "sloppy VaR" add-ons, which are based on the (wrong) assumption that estimation uncertainty is independent of the true ρ^*. For all considered sample sizes (N, T) and values of the point estimate, the former add-ons are roughly twice the size of the latter.

That said, when the point estimate of correlation increases, the VaR add-on induced by noise in this estimate decreases as a share of the naive VaR. Continuing with the above example, in which $N = 200$ and $T = 120$ months, the add-on decreases from 3.4% of the naive VaR when $\widehat{\rho} = 10\%$ to 2.2% of the naive VaR when $\widehat{\rho} = 30\%$. The flipside of this result is that the naive VaR, which abstracts from the noise in parameter estimates and, thus, overstates their information content, is more sensitive to changes in these estimates than the correct VaR (i.e. the naive VaR plus the add-on).

The underlying framework allows for a straightforward derivation of VaR^{alt} add-ons that focus directly on specific quantiles of the distribution of ρ^c and the common credit-risk factor, M (recall equation (5)). The two right-most panels of Table 2 quantify two versions of these add-ons. One of them arises when the dependence of estimation noise on the true parameter value is taken into account ("alternative add-on") and the other one when it is not ("alternative sloppy add-on"). As anticipated in Section 1.2, the alternative add-ons, ranging between 35% and 71% of the corresponding naive VaR, are substantially larger than the correct add-ons.

3.2 PD uncertainty

When one considers solely uncertainty about the probability of default, the analytic challenges are the reverse of these encountered in the previous subsection. Namely, the Cramer-Rao lower bound on the variance of the noise in PD estimates does not exist in closed form but, for a given value of this lower bound, the VaR measure itself does. In order to see the latter implication, let the true correlation be observed without noise $(\widehat{\rho} = \rho^*)$ and recall that the distribution of the common credit risk factor, M, and the posterior distribution of the default point, $\Phi^{-1}(PD^c)$, are both normal. By expression (2), this means that the investor effectively faces a single common risk factor, which equals $\Phi^{-1}(PD^c) - \sqrt{\widehat{\rho}}M$ and has a normal distribution. As a result, the ASRF model is applicable, implying that the portfolio VaR at the $(1-\alpha)$ confidence level equals:[17]

$$VaR\left(\widehat{PD},\widehat{\rho}\right)=\Phi\left(\frac{\mu_D\left(\widehat{PD},\widehat{\rho}\right)-\sqrt{\widehat{\rho}+\sigma_D^2\left(\widehat{PD},\widehat{\rho}\right)}\Phi^{-1}(\alpha)}{\sqrt{1-\widehat{\rho}}}\right) \tag{10}$$

where $\mu_D\left(\widehat{PD},\widehat{\rho}\right)$ and $\sigma_D^2\left(\widehat{PD},\widehat{\rho}\right)$ denote the mean and variance of the posterior distribution of the default point. Expression (10) leads to VaR add-ons that are reported in the left-hand panels of Table 1 as well as in Table 3.

A comparison between the left-hand and middle panels of Table 1 points to three important differences between the impact of PD uncertainty on the VaR perceived by the investor and the corresponding impact of correlation uncertainty. First, the VaR add-on induced by uncertainty in the PD estimate is much higher. If, for example, $\rho^*=20\%$ and the point estimate $\widehat{PD}=1\%$ is obtained from data covering $N=200$ obligors over $T=10$ years, the correct add-on is 3.17 cents on the dollar, or 22% of the naive VaR. This add-on is almost 8 times larger than the corresponding add-on induced by correlation uncertainty.

The difference between these add-ons is a natural consequence of the fact that the uncertainty about the probability of default is greater than the uncertainty about the asset-return correlation. In turn, the relative size of the two types of uncertainty is mostly driven by the maintained assumption that there are fewer data points for default rates than for asset returns (see Section 2.1 above).[18] For an implication of this assumption, suppose that the true parameter values are $\rho^*=20\%$ and $PD^*=1\%$, and the sample is of size $N=200$, $T=10$. In this case, the Cramer-Rao lower bound on the standard deviation of the PD estimate is 0.47%, i.e. almost half of the true parameter value (see the second panel of Table 4). By contrast, the Cramer-Rao lower bound on the standard deviation of the correlation estimate is 2.1%, which is slightly more than one-tenth of the true parameter value (see Figure 2).

The second noteworthy difference from the case of correlation uncertainty is that the VaR add-on induced by PD uncertainty is more sensitive to changes in the size of the cross section. For example, when $\rho^*=20\%$, $\widehat{PD}=1\%$ and $T=10$, reducing the cross section of the dataset from $N=200$ to 50 obligors increases the VaR add-on by 60% (from 3.17 to 5.07 cents on the dollar). By contrast, as seen above, the corresponding change is virtually nil in the context of correlation uncertainty. Furthermore, raising

[17] This expression incorporates the independence between PD^c and M.

[18] Another important assumption in this context is that asset returns are observed directly. This assumption is relaxed in Section 4.2 below.

$N = 200$ to 1000 obligors depresses the VaR add-on by a non-negligible 21%. That said, at 2.49 cents on the dollar or 17% of the naive VaR, the VaR add-on remains substantial even when $N = 1000$, i.e. when there are data on all the exposures in a virtually asymptotic portfolio.[19]

Third, analysis of PD uncertainty provides the only examples in which a rise in the point estimate can lead to a decline in the correct add-on, together with a rise in the naive VaR. If, for example, $N = 50$ and the time series covers $T = 5$ years, a rise in the point estimate from $\widehat{PD} = 1\%$ to $\widehat{PD} = 5\%$ leads to a decline in the add-on from 10.77 to 8.92 cents on the dollar (see the left-hand panels of Table 1). This is another manifestation of the fact that, since it accounts for noise in parameter estimates, the correct VaR is less sensitive to changes in these estimates than the naive VaR, which abstracts from estimation noise. The difference between the alternative sensitivities is greater in the context of PD uncertainty because this uncertainty is greater than that about the asset-return correlation.

Besides these three differences, there are a number of qualitative similarities between the effect of PD uncertainty and that of correlation uncertainty on VaR measures. In particular, a comparison between Tables 2 and 3 reveals similar consequences of: (i) changing the time span of the data, (ii) ignoring the dependence of estimation noise on the true parameter, (iii) focusing directly on specific quantiles of the credit- and estimation-risk factors.

3.3 Combining the two sources of uncertainty

The interaction between noise in the estimator of the asset-return correlation and noise in the PD estimator inflates the VaR perceived by the investor. In order to see what drives this result, it is useful to step back and recall the sequential estimation procedure. As outlined in Section 2.4, the investor first derives the posterior distribution of candidate values for the correlation coefficient. Then, conditioning on such a candidate value and the observed default rates, the investor obtains a point estimate of the PD.

For the considered empirical framework and parameter values, a higher correlation candidate induces a higher point estimate of the PD. This is illustrated in Figure 3, where the squares plot the typical distribution of 10 one-year default rates for a true correlation $\rho^* = 20\%$ and a true $PD^* = 1\%$.[20] In addition, the blue stars plot the

[19] Tarashev and Zhu (2008a) demonstrate that a portfolio of 1000 homogeneous exposures can be safely treated as perfectly fine grained (or asymptotic).

[20] This typical distribution is based on 100,000 sets of 10 one-year default rates, which are simulated for the given PD^* and ρ^*. Then, the ten default rates in each set are ordered from lowest to highest. In turn, ordered sets are used as columns in a $10 \times 100,000$ matrix of default rates. The median of the

distribution of joint defaults implied by a low value of the correlation candidate, denoted by ρ_1^c in the figure, and the corresponding estimate of the probability of default, $PD\,(\rho_1^c)$.[21] In turn, the red stars plot the probability distribution of joint defaults obtained by combining the same $PD\,(\rho_1^c)$ with a higher correlation candidate, ρ_2^c. The differences between the blue and red stars arise from the fact that, all else constant, a higher correlation increases the probability of default clustering, which is manifested in a higher probability of zero and many defaults but lower probabilities of intermediate numbers of defaults. Importantly, since the initial pair $\{\rho_1^c, PD\,(\rho_1^c)\}$ delivers an optimal fit to the typical default rates (blue stars), using $PD\,(\rho_1^c)$ with ρ_2^c worsens this fit (red stars). By extension, the point estimate $PD\,(\rho_2^c)$, which is based on the higher correlation candidate ρ_2^c, leads to a reversal of this worsening of the fit (green stars). Since $PD\,(\rho_2^c) > PD\,(\rho_1^c)$ the reversal involves a lower (higher) probability of zero (a positive number of) defaults.

The positive relationship between the correlation candidate and the PD point estimate imputes a positive correlation in the joint posterior distribution of the candidate values ρ^c and PD^c. In turn, this positive correlation increases the VaR measure, albeit only slightly (refer to Table 1). In the benchmark example, where the point estimates are $\widehat{PD} = 1\%$, $\widehat{\rho} = 20\%$ and there are $T = 10$ years of data on $N = 200$ obligors, the correct VaR add-on (3.93 cents on the dollar) is 10% larger than the sum of the two add-ons associated, respectively, with noise only in the PD or correlation estimates ($3.17 + 0.41$ cents on the dollar).

3.4 A closed-form approximation of the VaR measure

An important by-product of the analysis is that the closed-form approximate measure in expression (10), which incorporates only uncertainty about the PD, goes a long away in accounting for the VaR perceived by the investor. This is clearly seen by referring to Table 1 and reconstructing VaR measures by adding the add-ons reported in the

default rates in the first row of this matrix is plotted in Figure 3 as the lowest value in the typical draw of default rates. The median in the second row is plotted as the second lowest value in the typical draw of default rates, etc.

[21] More precisely, an implied probability distribution attributes non-zero probability mass to each default rate in the set $\{0, 1/N, 2/N, \cdots 1\}$, which in the specific example equals $\{0, 0.005, 0.01, \cdots 1\}$. In order to facilitate comparisons with the typical distribution of default rates (squares in Figure 3), an implied distribution (stars) is plotted as follows. The implied probabilities of the default rate being equal to 0 or 0.005 are plotted directly. The implied probability of the default rate being equal to 0.01 and half of the implied probability of the default rate being equal to 0.015 are added and their sum equals the height of a star at 0.01. Half of the implied probability of the default rate being equal to 0.015 and the probabilities of the default rate being equal to 0.02 and 0.025 are added and their sum equals the height of a star at 0.02. The implied probabilities of the default rate being equal to 0.03 or more are added and their sum equals the height of a star at 0.035.

left- and right-hand panels to the corresponding naive VaRs. For the point estimates and sample sizes considered for this table, VaR measures that incorporate only PD uncertainty tend to understate the corresponding correct VaRs by 5% or less.

Importantly, the closed-form expression in (10) has two general computational advantages. First, it reduces the computational burden by limiting the Bayesian inference procedure to a single parameter: the probability of default. Second, for a given posterior distribution of PD candidate values, this expression delivers a VaR measure directly, without relying on numerical simulations.

4 Extensions

This section considers two extensions of the estimation procedure, which address aspects of the investor's information set that the analysis has so far abstracted from. The first subsection illustrates how richer prior beliefs about credit-risk parameters can affect the VaR measure. Then, the second subsection shows that noise in observed asset returns complicates the inference procedure and then proposes how to account for such noise in a computationally efficient way.

4.1 Prior beliefs about probability of default

The results reported in Section 3 are based on diffuse priors regarding the asset-return correlation and the PD. The benefit of assuming such priors is that they render transparent the transition from the conditional PDF of the estimator to the posterior PDF of the parameter (see equations (6) and (7)). That said, an investor's information about a parameter may go beyond the information contained in the dataset that he/she uses in order to obtain a point estimate of this parameter. In other words, the prior belief may not be diffuse, which would affect the parameter's posterior distribution and, consequently, the perceived VaR in important ways.

Consider an investor who has a short forecast horizon and is interested in the VaR of a portfolio of homogeneous obligors that have a particular credit rating. Suppose further that the investor observes a rather long time series of the one-year default rates in the same rating class. Given the investor's short horizon and assuming rather frequent, albeit persistent, changes in the (average) PD within the rating class, the investor might derive a posterior distribution of PD candidates that relies more heavily on recent default rates. In the light of equation (7), one way of doing this is to let all of the observed default rates determine prior beliefs but to derive a point estimate of the PD only on the basis of the last several default rates.

The extent to which prior beliefs affect the VaR measure depends on the extent to which they are in accordance with the point estimate. To take a concrete example, suppose that the investor bases his/her prior on the one-year default-rates of all corporate obligors rated BB by Moody's from 1990 to 2007. The average of these default rates is 0.98% and their standard deviation is 1.18%.[22] Suppose further that the investor obtains a point estimate $\widehat{PD} = 1\%$ on the basis of default data covering $N = 200$ BB-rated obligors over $T = 5$ years and, for simplicity, observes directly the true correlation $\rho^* = 20\%$. Since, in this example, the prior mean and the point estimate are quite close, the main effect of the prior is to tighten the posterior distribution of PD candidates relative to that implied by a diffuse prior. Being tantamount to less parameter uncertainty, this results in a VaR add-on that equals 3.25 cents on the dollar, down from 6.41 cents under a diffuse prior (refer to Table 3).

Importantly, the result could be quite different if, keeping all else constant, the investor's dataset covered B-rated corporate obligors. A point estimate $\widehat{PD} = 1\%$ for B-rated obligors is admittedly extreme but not unreasonable, given that the default rate of such obligors has averaged 1.01% between 2003 and 2007. However, such a point estimate would be quite at odds with a prior belief that incorporates the default history of B-rated corporate obligors since 1990: this history features an average default rate of 5% and a standard deviation of default rates of 4%. Against such a prior, the point estimate appears overly optimistic. Consequently, the implied posterior distribution attributes more probability mass to high PD candidates than a posterior based on a diffuse prior. Not surprisingly then, the end result is a VaR add-on of 9.15 cents on the dollar, up from 6.41 cents under a diffuse prior.

The above two illustrative examples indicate that departures from the assumption of diffuse priors can lead to substantial changes in perceived VaR. These changes may stem from an improved precision of the information set, as most clearly seen in the example with BB-rated obligors, or from an alignment of posterior beliefs with long-term default experience, as seen in the example with B-rated obligors. That said, a rigorous analysis of prior distributions and their transformation into posterior beliefs would have to be based explicitly on a learning process, which is beyond the scope of this paper.

4.2 Noise in observed asset returns

In real-life applications, asset returns – or, more generally, the stochastic drivers of default events – would be observed with noise. Depending on whether the noise is idiosyncratic (driven by obligor-specific imperfections in the measurement of assets'

[22] The data source is the Credit Risk Calculator database of Moody's Investors Service.

market value) or systematic (a result, for example, of mapping equity prices of different obligors into corresponding asset values via the same mis-specified model), it could lead to a downward or upward bias in correlation estimates.

To see why, generalize (1) to account for observation noise, which is denoted by $U_{i\tau}$:

$$\begin{aligned}
\ln\left(V_{i,\tau}\right)-\ln\left(V_{i,\tau-1}\right) &= \mu^*\Delta+\sigma^*\sqrt{\Delta}\sqrt{1-\psi^*}\left(\sqrt{\rho^*}M_\tau+\sqrt{1-\rho^*}Z_{i,\tau}\right) \\
&\quad +\sigma^*\sqrt{\Delta}\sqrt{\psi^*}U_{it} \\
U_{i\tau} &= \sqrt{\lambda^*}M_\tau^U+\sqrt{1-\lambda^*}Z_{i\tau}^U \\
M_\tau^U &\sim N\left(0,1\right) \text{ is independent of } Z_{i\tau}^U\sim N\left(0,1\right) \\
\psi^* &\in \left[0,1\right],\ \lambda^*\in\left[0,1\right]
\end{aligned}$$

where ψ^* and λ^* control, respectively, the amount of observation noise and the systematic component of this noise. Then note that the true correlation of observed asset returns equals

$$\rho^{*,OBS}=\left(1-\psi^*\right)\rho^*+\lambda^*\psi^* \tag{11}$$

where ρ^* continues to indicate the true correlation of actual asset returns. All else constant, $\rho^{*,OBS}$ increases in the systematic component of the observation noise, λ^*. The level of $\rho^{*,OBS}$ also increases in the overall amount of noise, ψ^*, if the systematic component is large enough, i.e. if $\lambda^*>\rho^*$.

An important implication of (11) is that, if the data comprise only noisy observations of asset returns, the true correlation of actual asset returns is unidentifiable. Namely, for any given $\rho^{*,OBS}$, one can pick any admissible value of the correlation of actual asset returns, ρ^*, and find a continuum of different pairs (λ^*,ψ^*) that render $\rho^{*,OBS}$ and ρ^* mutually consistent. This is illustrated in Figure 4.

Given that the asset-return correlation is unidentifiable, an investor could impose a conservative upward adjustment on the point estimate $\widehat{\rho}$ and treat the adjusted value as if it were coming from data that are free of observation noise. By equation (11), a possible conservative adjustment would be one consistent with no systematic observation noise ($\lambda^*=0$) and some idiosyncratic observation noise ($\psi^*>0$). As illustrated by Table 5, such an approach results in significant upward revisions of the VaR measure. Suppose, for example, that the true probability of default $PD^*=1\%$ is known and the correlation is estimated at $\widehat{\rho}=20\%$ on the basis of data that cover $N=200$ obligors over $T=120$ months. Then, setting $\psi^*=0.05$ (i.e. allowing idiosyncratic observation noise to accounts for 5% of the variability of observed asset returns) leads to a VaR add-on that is almost three times as high as the add-on obtained for $\psi^*=0$. However, in order to match the add-on induced by PD uncertainty when $N = 200$, $T = 10$ years, $\widehat{PD} = 1\%$ and the correlation is known to be $\rho^* = 20\%$ (recall Table 3), it is necessary to set $\psi^* = 0.15$.

In order to incorporate asset-return observation noise in the closed-form VaR approximation (10), it is necessary to abstract from the estimation noise studied above in Section 3.1. In other words, it is necessary to treat the point estimate $\widehat{\rho}$ as if it were equal to $\rho^{*,OBS}$. Then, equation (11) allows to map $\rho^{*,OBS}$ into a correlation of actual asset returns, ρ^*, which can be used directly in (10). Of course, the mapping and, thus, the VaR measure will be affected by the parameterization of the idiosyncratic and systematic observation noise. The effect of different parameterizations is illustrated in the "$T = \infty$" rows in Table 5, which contain VaR add-ons that are based on infinite time series of data and, thus, on perfect knowledge of $\rho^{*,OBS}$.

The discussion in this section has so far abstracted from data on default rates. In principle, this is important because, when asset returns are observed with noise, such data do provide useful information about the asset-return correlation. In the context of the adopted empirical framework, however, using data on default rates in order to make inference about the asset-return correlation would be of little value but would be associated with substantial computational burden. Here is why.

First, the correlation estimate based on default data is extremely imprecise, especially if the obligors are of moderate to high credit quality. For example, given a true $PD^* = 1\%$ and a true correlation $\rho^* = 20\%$, default data covering $N = 200$ obligors over $T = 10$ years imply that the noise in the most efficient unbiased estimator of ρ^* has a standard deviation of 12.5 percentage points. To put this into perspective, note that: (i) realistic values of the (average) asset-return correlation are between 5% and 45% and (ii) the standard deviation of a uniform random variable with support from 5% to 45% equals 11.55%.

Second, a VaR measure that incorporates explicitly inference about observation noise in asset returns would face the so-called "curse of dimensionality". To see why, recall that the VaR measures discussed in Section 3 are based on posterior distributions that condition on two parameter estimates: $\widehat{\rho}$ and $\widehat{PD}$. In the presence of observation noise, properly constructed posterior distributions would need to condition on two more estimates: those of the observation noise parameters ψ^* and λ^*.[23] Since the number of numerical simulations grows exponentially in the number of parameters that the inference procedure is applied to, doubling the latter number would be prohibitively burdensome.

[23] This is because the amount of estimation noise in correlation and PD estimates depends on the values of the noise parameters φ^* and ψ^*.

Conclusion

This paper has analyzed the impact of parameter uncertainty on the VaR perceived by an investor in a homogeneous asymptotic credit portfolio. The main conclusion is that this impact is strong for a wide range of portfolio characteristics and for a wide range of dataset sizes. As a useful by-product, the analysis has delivered an approximate VaR measure, which exists in closed-form and is, thus, computationally convenient. This measure accounts for PD uncertainty and, given a judicious adjustment for noise in observed asset returns, could approximate well the correct VaR measure.

Relaxing some of the assumptions adopted by this paper would provide for fruitful directions of future research. One is to address rigorously the issue of parameter heterogeneity in the context of credit VaR measures. Given that deriving and simulating the joint posterior distribution of a large number of heterogeneous parameters is likely to impose an insurmountable computational burden, it is important to establish conditions under which it is justifiable to focus solely on noise in the estimator of a representative (e.g. the average) parameter. Another possible direction of future research is to consider cyclical developments in credit conditions, which make credit-risk parameters change over time and, consequently, impair the estimates of these parameters.

Appendix A

In order to render the appendix self-contained, the statement of Proposition 1 contains more information here than in the main text.

Proposition 1: Denote the correct VaR of the portfolio at the $(1-\alpha)$ confidence level by $Q_{1-\alpha}(L_n)$. Define $X \equiv \{M, PD^c, \rho^c\}$. The expectation $E(L_n|X)$ does not depend on n and can be denoted by $\tilde{\Phi}(X)$, where $\tilde{\Phi}(\cdot)$ is an analytic function. Denote the $(1-\alpha)$ quantile of the distribution of $\tilde{\Phi}(X)$ by $Q_{1-\alpha}\left(\tilde{\Phi}(X)\right)$. As $n \to \infty$, $Q_{1-\alpha}(L_n) - Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) \to 0$.

Proof. The proof relies on three aspects of the model:

Aspect 1: The assumed homogeneity of parameters across exposures implies that the following holds trivially:

$$\lim_{n\to\infty} E(L_n|X) = E(L|X) = \Phi\left(\frac{\Phi^{-1}(PD^c) - \sqrt{\rho^c}M}{\sqrt{1-\rho^c}}\right) \equiv \tilde{\Phi}(X) \tag{12}$$

Aspect 2: The expected loss $E(L|X=x) \equiv \tilde{\Phi}(x)$ changes continuously in x.

Aspect 3: The PDF of X is well-defined and continuous everywhere on its support.

By the law of iterated expectations, it follows that

$$\lim_{n\to\infty} \Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right)\right) = \lim_{n\to\infty} \int_{\mathbf{R}^3} \Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) | X = x\right) dF(x) \tag{13}$$

where $F(x)$ is the CDF of X.

In addition, Proposition 1 in Gordy (2003), which is applicable owing to Aspects 1-3 above, implies that $L_n|x \to \tilde{\Phi}(x)$ almost surely as $n \to \infty$.

Given this implication, the fact that $F(\cdot)$ is absolutely continuous (by Aspects 2 and 3) and that $\Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) |X = \cdot\right)$ is bounded between 0 and 1, the dominated convergence theorem applies (see Billingsley (1995)). Thus:

$$\begin{aligned} & \lim_{n\to\infty} \int_{\mathbf{R}^3} \Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) | X = x\right) dF(x) \\ = & \int_{\mathbf{R}^3} \lim_{n\to\infty} \Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) | X = x\right) dF(x) \\ = & \int_{\mathbf{R}^3} \Pr\left(\tilde{\Phi}(x) \leq Q_{1-\alpha}\left(\tilde{\Phi}(x)\right)\right) dF(x) \\ = & \Pr\left(\tilde{\Phi}(X) \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right)\right) = 1 - \alpha \end{aligned} \tag{14}$$

where the second equality follows from Proposition 1 in Gordy (2003). The third equality in expression (14) follows from (12). In turn, the fourth equality is a result of the fact that $\Pr\left(\tilde{\Phi}(x) \leq \cdot\right)$ is continuous.

Combining expressions (13) and (14), it follows that:

$$\lim_{n\to\infty} Q_{1-\alpha}(L_n) \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right)$$

Repeating the steps from expression (13), it follows that, for any $\varepsilon > 0$, there exists $\delta > 0$ such that:

$$\lim_{n\to\infty} \Pr\left(L_n \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) - \varepsilon\right) = (1-\alpha) - \delta$$

Thus,

$$Q_{1-\alpha}\left(\tilde{\Phi}(X)\right) - \varepsilon \leq \lim_{n\to\infty} Q_{1-\alpha}(L_n) \leq Q_{1-\alpha}\left(\tilde{\Phi}(X)\right)$$

Since ε can be arbitrarily close to 0, the proof is complete. ■

Appendix B

This appendix derives the Cramer-Rao lower bounds on the standard deviations of the noise of correlation and PD estimates. These lower bounds are denoted by $\sigma_\rho(\rho^c)$ and $\sigma_{PD}(PD^c, \rho^c)$ in the main text.

Given that the data generating process is as specified in (1) and (2) and the dataset is as described in Section 2.1, the log likelihood of asset returns on a particular date $t \in \{1, 2, \cdots, T\}$ is simply the log-likelihood of N jointly-normal random variables, with mean μ^*, variance σ^{*2} and correlation ρ^c. Denote this log-likelihood by $LL^{ar}\left(\{V_{it}\}_{i=1}^N; \theta^*\right)$, where $\theta^* \equiv \{\mu^*, \sigma^*, \rho^c\}$. The Fischer information matrix is then given by

$$I(\theta^*) = E\left(\left.\frac{\partial^2 LL^{ar}\left(\{V_{it}\}_{i=1}^N; \theta\right)}{\partial\theta^2}\right|_{\theta=\theta^*}\right)$$

Given that asset returns are serially uncorrelated, the Cramer-Rao lower bounds on the variances of parameter estimators are as reported in the following matrix

$$\frac{1}{T}I^{-1}(\theta^*) = \begin{matrix} \frac{(\sigma^*)^2(1+(N-1)\rho^c)}{NT} & 0 & 0 \\ 0 & \frac{2(\sigma^*)^4\left(1+(N-1)(\rho^c)^2\right)}{NT} & \frac{2(\sigma^*)^2\rho^c(1-\rho^c)(1+(N-1)\rho^c)}{NT} \\ 0 & \frac{2(\sigma^*)^2\rho^c(1-\rho^c)(1+(N-1)\rho^c)}{NT} & \frac{2(1-\rho^c)^2(1+(N-1)\rho^c)^2}{TN(N-1)} \end{matrix}$$

The $(3,3)$ element of this matrix refers to the variance of the correlation estimator:

$$\sigma_\rho^2(\rho^c) = \frac{2(1-\rho^c)^2(1+(N-1)\rho^c)^2}{TN(N-1)}$$

This Cramer-Rao lower bound is:

1. not affected neither by the value of μ^* nor by the fact that this value has to be estimated;

2. not affected by the value of σ^* but inflated by the fact that this value has to be estimated;

3. declines to 0 as the length of the time series, T, increases;

4. declines to $2(\rho^c(1-\rho^c))^2 \Big/ T > 0$ as the size of the of the cross-section increases;

5. increases in ρ^c for $\rho^c \in \left(0, \frac{N-2}{2(N-1)}\right)$ but decreases in ρ^c for $\rho^c \in \left(\frac{N-2}{2(N-1)}, 1\right)$.

The date-t log-likelihood of defaults conditions on a candidate value of the correlation, ρ^c:

$$\begin{aligned} & LL^{dr}\left(\{d_{it}\}_{i=1}^{N};PD^c,\rho^c\right) \qquad (15)\\ = \; & \log\int\Phi\left(\frac{\Phi^{-1}\left(PD^c\right)-\sqrt{\rho^c}M}{\sqrt{1-\rho^c}}\right)^{D_t}\Phi\left(\frac{\sqrt{\rho^c}M-\Phi^{-1}\left(PD^c\right)}{\sqrt{1-\rho^c}}\right)^{N-D_t}\phi\left(M\right)dM \end{aligned}$$

where $d_i = 1$ if obligor i defaults and $d_i = 0$ otherwise; $D_t = \sum_{i=1}^{N} d_{it}$. Then, the Cramer-Rao lower bound on the noise in the estimate of PD^c is given by

$$\sigma^2_{PD}\left(PD^c,\rho^c\right) = 1\Big/T\cdot E\left(\left.\frac{d^2LL^{dr}\left(\{d_{it}\}_{i=1}^{N};\pi,\rho^c\right)}{d\pi^2}\right|_{\pi=PD^c}\right)$$

This lower bound is derived numerically.

As implied by (15), the Cramer-Rao lower bound on the variance of the noise in the PD estimate is unaffected neither by the value of μ^* nor by the fact that this value has to be estimated. Further, this Cramer-Rao lower bound does not depend on the value of σ^* but, via the posterior distribution of ρ^c, does depend on the fact that this value has to be estimated.

Appendix C

A discretization of the posterior distribution $h\left(\rho^c|\widehat{\rho}\right)$ has two effects. First, it reduces the number of Monte Carlo simulations necessary for an estimate of the distribution of the conditional loss, $\Phi\left(\Phi^{-1}\left(PD^c\right)-\sqrt{\rho^c}M\big/\sqrt{1-\rho^c}\right)$. Second, a discretization impairs the precision of this such an estimate. In order to address this trade-off, I proceed as follows:

1. Divide the support of ρ^c into a number of intervals, such that the probability mass associated with any interval is a scalar multiple of the largest probability mass associated with a single interval.

2. Focusing on a particular interval, I assign the entire associated probability mass to a single point that equals the expected value of ρ^c, conditional on ρ^c belonging to the interval.

3. I increase the number of intervals and repeat steps 1. and 2. This process continues until the implied relative change in portfolio VaR is less than 1%.

The final discretization adopted by this paper assigns probability mass to 10 points on the support of ρ^c. Specifically, probability of 2.5% is assigned to 4 points that equal the expected value of ρ^c, conditional on $\rho^c \in [-\infty, 2.5\%], [2.5\%, 5\%], [95\%, 97.5\%]$ or $[97.5\%, +\infty]$. In addition, probability mass of 15% is assigned to 6 six points that equal the expected value of ρ^c, conditional on
$\rho^c \in [5\%, 20\%], [20\%, 35\%], [35\%, 50\%], [50\%, 65\%], [65\%, 80\%]$ or $[80\%, 95\%]$.

Thus, the discretized version of the posterior $h(\rho^c|\widehat{\rho})$ can be simulated with as a few as 40 draws of ρ^c: 1 draw of each of the two smallest and two largest values on the discretized support and 6 draws of each of the six intermediate values.

References

Billingsley, P. 1995. Probability and Measure. 3rd edition. Wiley. New York.

Borio, C. and K. Tsatsaronis. 2004. “Accounting and Prudential regulation: From Uncomfortable Bedfellows to Perfect Partners?”, Journal of Financial Stability 1: 111-135.

Cantor, R., D. Hamilton and J. Tennant. 2008. “Confidence Intervals for Corporate Default Rates”. Risk. March. 93-98.

Gössl, C. 2005. “Predictions Based on Certain Uncertainties – a Bayesian Credit Portfolio Approach.” Working paper.

Gordy, M. B. 2003. “A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules.” Journal of Financial Intermediation 12 (3): 199–232.

Gordy, M., and E. Lütkebohmert. 2007. “Granularity Adjustment for Basel II.” Deutsche Bundesbank Discussion Paper No. 01/2007.

Hanson, S. and T. Schuermann. 2006. “Confidence Intervals for Probabilities of Default.” Journal of Banking and Finance 30. 2281-2301.

Heitfield, E. 2008. “Parameter Uncertainty and the Credit Risk of Collateralized Debt Obligations.” Federal Reserve Board Working Paper.

Kupiec, P. 2008. “A Generalized Single Common Factor Model of Portfolio Credit Risk.” Journal of Derivatives 15 (Spring): 25–40.

Lando, D. and T. Skodeberg. 2002. “Analyzing Rating Transitions and Rating Drift with Continuous Observations”. Journal of Banking and Finance 26. 423-444.

Löffler, G. 2003. “The Effects of Estimation Error on Measures of Portfolio Credit Risk.” Journal of Banking and Finance 27 (8): 1427–53.

Tarashev, N., and H. Zhu. 2008a. “Specification and Calibration Errors in Measures of Portfolio Credit Risk.” International Journal of Central Banking, vol 4, no 2, June.

Tarashev, N., and H. Zhu. 2008b. “The Pricing of Correlated Default Risk: Evidence from the Credit Derivatives Market.” Journal of Fixed Income, Summer.

T A B L E – 1 –

Impact of estimation noise on portfolio VaR (in per cent)

$\widehat{PD}$ = 1%; $\hat{\rho}$ = 20%										
	Naïve VaR	Noise in $\widehat{PD}$ only; add-ons to naïve VaR			Noise in $\hat{\rho}$ only; add-ons to naïve VaR			Noise in both $\widehat{PD}$ and $\hat{\rho}$: add-ons to naïve VaR		
		N=50	200	1000	N=50	200	1000	N=50	200	1000
T=5 years (=60 months)	**14.55**	**10.77**	**6.41**	**4.98**	**0.91**	**0.82**	**0.81**	**13.10**	**8.20**	**7.00**
T=10 years (=120 months)		**5.07**	**3.17**	**2.49**	**0.42**	**0.41**	**0.40**	**5.85**	**3.93**	**3.22**

$\widehat{PD}$ = 5%; $\hat{\rho}$ = 20%										
	Naïve VaR	Noise in $\widehat{PD}$ only; add-ons to naïve VaR			Noise in $\hat{\rho}$ only; add-ons to naïve VaR			Noise in both $\widehat{PD}$ and $\hat{\rho}$: add-ons to naïve VaR		
		N=50	200	1000	N=50	200	1000	N=50	200	1000
T=5 years (=60 months)	**38.44**	**8.92**	**6.71**	**5.17**	**1.94**	**1.79**	**1.77**	**11.37**	**9.08**	**7.36**
T=10 years (=120 months)		**4.71**	**3.71**	**2.90**	**0.97**	**0.89**	**0.88**	**5.97**	**4.93**	**3.96**

Note: Results refer to the 99.9% VaR of an asymptotic homogeneous portfolio. Panel headings refer to particular point estimates of the probability of default, $\widehat{PD}$ and asset-return correlation, $\hat{\rho}$. Row headings, T, refer to the length of the time series underpinning $\widehat{PD}$ (years) and $\hat{\rho}$ (months). Column headings, N, refer to the number of obligors in the dataset underpinning $\widehat{PD}$ and $\hat{\rho}$.

T A B L E – 2 –

Impact of noise in the correlation estimate on portfolio VaR (in per cent)

$\hat{\rho}$ **= 10%**; *PD** **= 1%**													
	Naïve	Sloppy add-on			Correct add-on			Alternative sloppy add-on			Alternative add-on		
		N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000
T=60 months		**0.27**	**0.21**	**0.19**	**0.66**	**0.56**	**0.52**	**4.65**	**3.99**	**3.82**	**5.53**	**4.95**	**4.79**
	7.75												
T=120 months		**0.14**	**0.11**	**0.10**	**0.32**	**0.26**	**0.25**	**3.11**	**2.69**	**2.58**	**3.53**	**3.16**	**3.06**

$\hat{\rho}$ **= 20%**; *PD** **= 1%**													
	Naïve	Sloppy add-on			Correct add-on			Alternative sloppy add-on			Alternative add-on		
		N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000
T=60 months		**0.40**	**0.34**	**0.34**	**0.91**	**0.82**	**0.81**	**8.59**	**7.97**	**7.81**	**9.79**	**9.31**	**9.18**
	14.55												
T=120 months		**0.22**	**0.19**	**0.18**	**0.42**	**0.41**	**0.40**	**5.78**	**5.38**	**5.27**	**6.36**	**6.03**	**5.94**

$\hat{\rho}$ **= 30%**; *PD** **= 1%**													
	Naïve	Sloppy add-on			Correct add-on			Alternative sloppy add-on			Alternative add-on		
		N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000
T=60 months		**0.55**	**0.44**	**0.43**	**0.95**	**0.95**	**0.95**	**12.38**	**11.81**	**11.66**	**13.19**	**12.80**	**12.69**
	22.44												
T=120 months		**0.32**	**0.25**	**0.24**	**0.50**	**0.50**	**0.50**	**8.38**	**8.01**	**7.91**	**8.83**	**8.54**	**8.47**

Note: Results refer to the 99.9% VaR of an asymptotic homogeneous portfolio. Panel headings refer to a particular point estimate of the asset-return correlation, $\hat{\rho}$, and a particular true value of the probability of default, *PD**. Row headings, T, and column headings, N, **(continues on the next page)**

(continues from the previous page) refer, respectively to the size of the time series and cross section underpinning $\hat{\rho}$. "Naïve" = a VaR measure that abstracts from estimation noise. Add-ons are defined as follows: "sloppy" = the difference between (i) a VaR measure that ignores the dependence of estimation noise on the value of the true parameter and (ii) the naive VaR measure; "correct" = the difference between (i) the correct VaR measure (under the assumption that the *PD** is observed directly) and (ii) the naive VaR measure; "alternative sloppy" = the difference between (i) an alternative VaR measure based on the 99.9^{th} quantile of a probability distribution of correlation candidates that ignores the dependence of estimation noise on the true value of the asset-return correlation and (ii) the naive VaR measure; "alternative" = the difference between (i) an alternative VaR measure based on the 99.9^{th} quantile of the true posterior distribution of correlation candidates and (ii) the naive VaR measure.

T A B L E – 3 –

Impact of noise in the PD estimate on portfolio VaR (in per cent)

ρ* = 20%; $\widehat{PD}$ = 1%													
	Naive	Sloppy add-on			Correct add-on			Alternative sloppy add-on			Alternative add-on		
		N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000
T=5 years		**5.56**	**3.34**	**2.48**	**10.77**	**6.41**	**4.98**	**31.57**	**23.78**	**20.09**	**39.07**	**28.07**	**23.43**
	14.55												
T=10 years		**2.92**	**1.73**	**1.29**	**5.07**	**3.17**	**2.49**	**22.04**	**16.40**	**13.91**	**25.11**	**18.40**	**15.40**

ρ* = 20%; $\widehat{PD}$ = 5%													
	Naive	Sloppy add-on			Correct add-on			Alternative sloppy add-on			Alternative add-on		
		N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000	N=50	N=200	N=1000
T=5 years		**6.38**	**4.82**	**4.25**	**8.92**	**6.71**	**5.17**	**30.77**	**27.02**	**25.48**	**32.52**	**27.96**	**24.91**
	38.44												
T=10 years		**3.31**	**2.48**	**2.18**	**4.71**	**3.71**	**2.90**	**22.62**	**19.68**	**18.49**	**23.76**	**20.63**	**18.50**

Note: Results refer to the 99.9% VaR of an asymptotic homogeneous portfolio. Panel headings refer to a particular point estimate of the probability of default, $\widehat{PD}$, and a particular true value of the asset-return correlation, ρ*. Row headings, T, and column headings, N, refer, respectively to the size of the time series and cross section underpinning $\widehat{PD}$. "Naïve" = a VaR measure that abstracts from estimation noise. Add-ons are defined as follows: "sloppy" = the difference between (i) a VaR measure that ignores the dependence of estimation noise on the value of the true parameter and (ii) the naive VaR measure; "correct" = the difference between (i) the correct VaR measure (under the assumption that ρ* is observed directly) and (ii) the naive VaR measure; "alternative sloppy" = the difference between (i) an alternative VaR measure based on the 99.9th quantile of a probability distribution of PD candidates that ignores the dependence of estimation noise on the true value of the probability of default and (ii) the naive VaR measure; "alternative" = the difference between an alternative VaR measure based on the 99.9th quantile of the true posterior distribution of PD candidates and (ii) the naive VaR measure.

T A B L E – 4 –

Errors in PD estimators

(Cramer-Rao lower bound on the standard deviation of unbiased estimators, in per cent)

***ρ** = 10%; *PD** = 1%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**0.76**	**0.51**	**0.41**
T = 10 years	**0.54**	**0.36**	**0.29**
T = 20 years	**0.38**	**0.26**	**0.21**
***ρ** = 20%; *PD** = 1%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**0.88**	**0.67**	**0.57**
T = 10 years	**0.62**	**0.47**	**0.40**
T = 20 years	**0.44**	**0.33**	**0.28**
***ρ** = 30%; *PD** = 1%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**1.02**	**0.81**	**0.69**
T = 10 years	**0.72**	**0.57**	**0.49**
T = 20 years	**0.51**	**0.41**	**0.35**
***ρ** = 10%; *PD** = 5%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**2.04**	**1.63**	**1.50**
T = 10 years	**1.44**	**1.15**	**1.06**
T = 20 years	**1.02**	**0.82**	**0.75**
***ρ** = 20%; *PD** = 5%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**2.56**	**2.20**	**2.06**
T = 10 years	**1.81**	**1.56**	**1.46**
T = 20 years	**1.28**	**1.10**	**1.03**
***ρ** = 30%; *PD** = 5%**			
	N = 50 obligors	N = 200 obligors	N = 1000 obligors
T = 5 years	**2.94**	**2.54**	**2.11**
T = 10 years	**2.08**	**1.79**	**1.49**
T = 20 years	**1.47**	**1.27**	**1.05**

Note: Panel headings indicate the true values of credit-risk parameters.

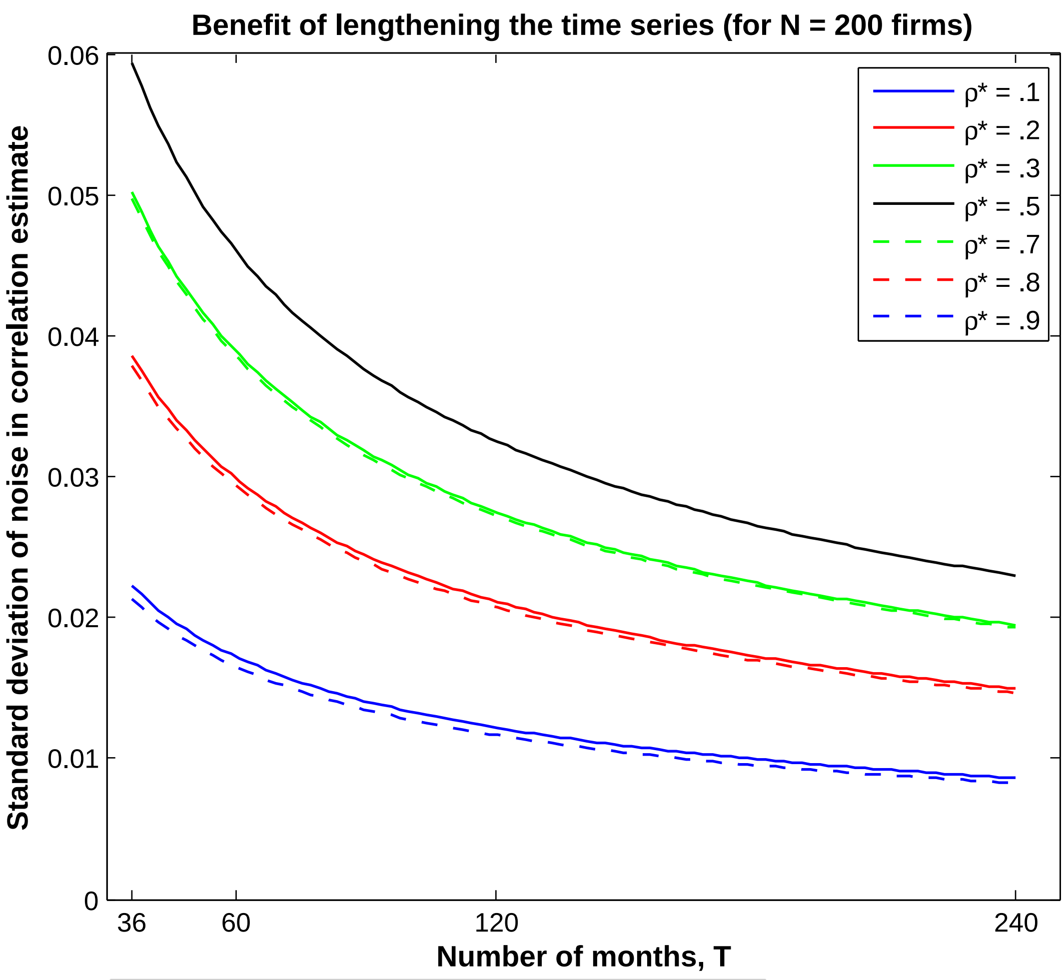
T A B L E – 5 –

Impact of noise in observed asset returns on portfolio VaR (in per cent)

$\hat{\rho}$ = 10%; *PD = 1%**											
	Naive	Correct add-on									
		N = 50 obligors					N = 200 obligors				
	7.75	ψ* = 0.0	ψ*= 0.05	ψ*= 0.1	ψ*= 0.15	ψ*= 0.20	ψ* = 0.00	ψ*= 0.05	ψ*= 0.10	ψ*= 0.15	ψ*= 0.20
T=60 months		**0.66**	**1.00**	**1.44**	**1.90**	**2.42**	**0.56**	**0.90**	**1.32**	**1.79**	**2.33**
T=120 months		**0.32**	**0.68**	**1.06**	**1.53**	**2.01**	**0.26**	**0.64**	**1.03**	**1.46**	**1.95**
T = ∞		**0**	**0.33**	**0.71**	**1.13**	**1.61**	**0**	**0.33**	**0.71**	**1.13**	**1.61**
$\hat{\rho}$ = 20%; *PD = 1%**											
	Naive	Correct add-on									
	14.55	N = 50 obligors					N = 200 obligors				
		ψ* = 0.0	ψ*= 0.05	ψ*= 0.1	ψ*= 0.15	ψ*= 0.20	ψ* = 0.00	ψ*= 0.05	ψ*= 0.10	ψ*= 0.15	ψ*= 0.20
T=60 months		**0.91**	**1.71**	**2.70**	**3.78**	**5.06**	**0.82**	**1.68**	**2.59**	**3.71**	**4.93**
T=120 months		**0.42**	**1.26**	**2.19**	**3.22**	**4.42**	**0.41**	**1.21**	**2.13**	**3.22**	**4.37**
T = ∞		**0**	**0.78**	**1.66**	**2.65**	**3.80**	**0**	**0.78**	**1.66**	**2.65**	**3.80**
$\hat{\rho}$ = 30%; *PD = 1%**											
	Naive	Correct add-on									
	22.44	N = 50 obligors					N = 200 obligors				
		ψ* = 0.0	ψ*= 0.05	ψ*= 0.1	ψ*= 0.15	ψ*= 0.20	ψ* = 0.00	ψ*= 0.05	ψ*= 0.10	ψ*= 0.15	ψ*= 0.20
T=60 months		**0.95**	**2.43**	**4.01**	**5.88**	**8.21**	**0.95**	**2.39**	**4.00**	**5.86**	**8.09**
T=120 months		**0.50**	**1.84**	**3.45**	**5.27**	**7.39**	**0.51**	**1.86**	**3.40**	**5.22**	**7.37**
T = ∞		**0**	**1.35**	**2.89**	**4.66**	**6.71**	**0**	**1.35**	**2.89**	**4.66**	**6.71**

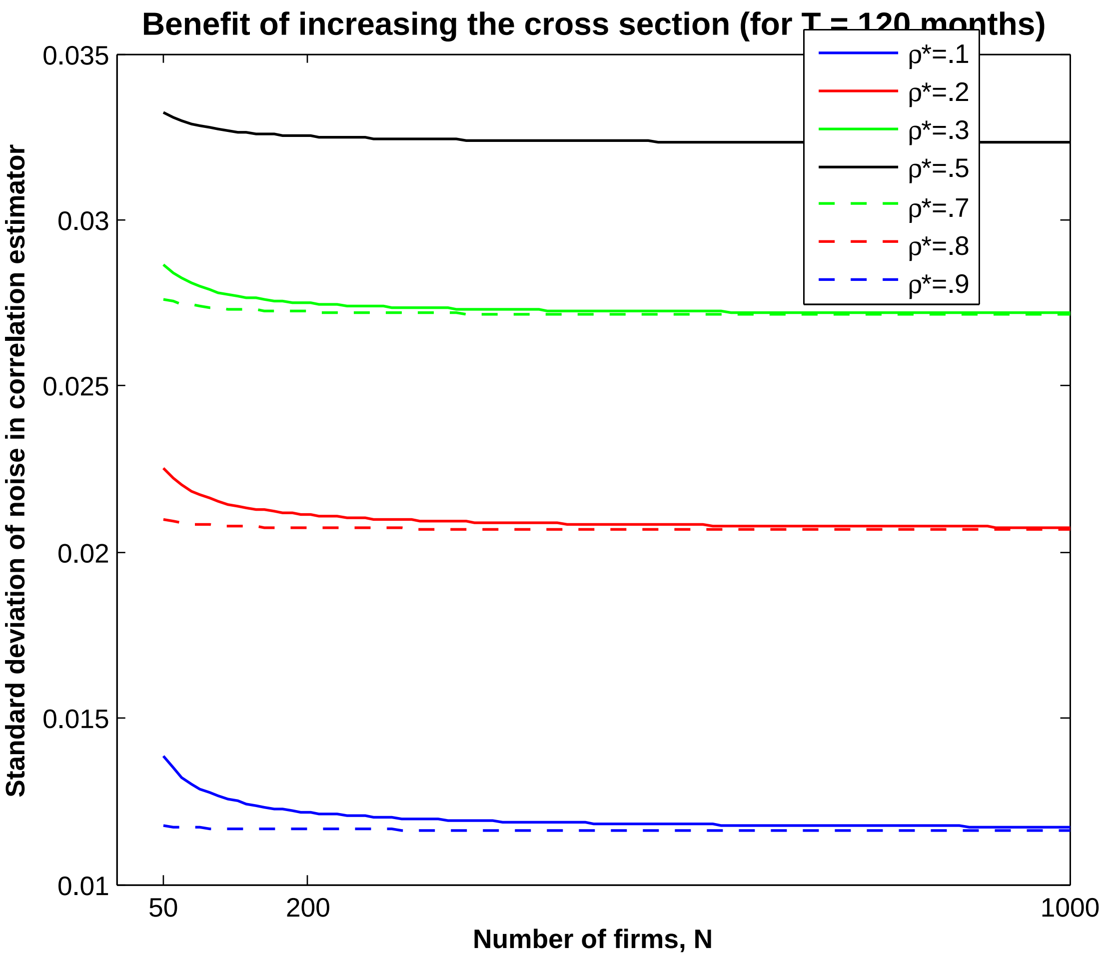
Note: The value of ψ* captures the amount of systemic noise in observed asset returns (refer to equation (11) and set λ* = 0). See Table 2 for further explanation.

Figure 1



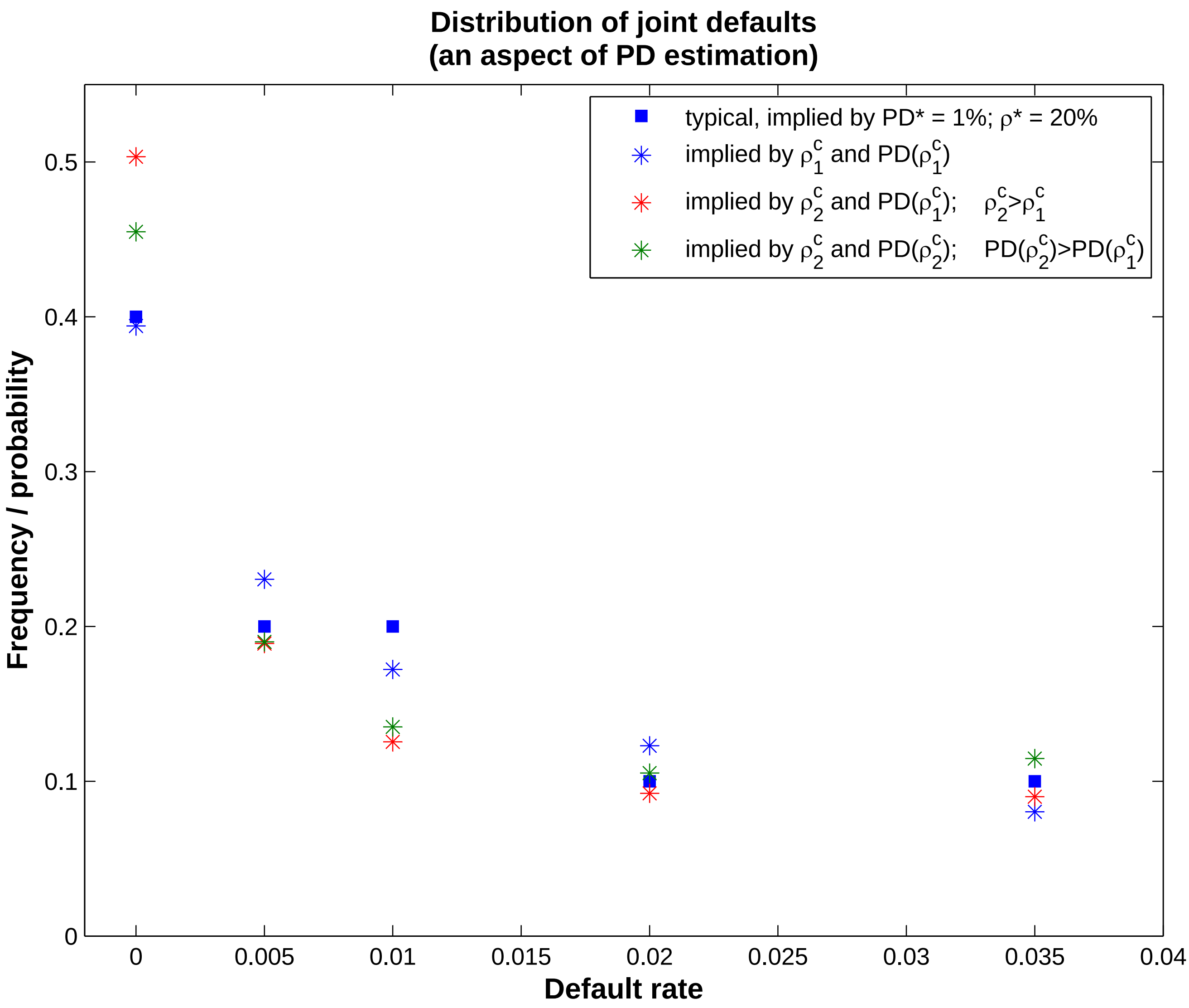
Note: All lines are plotted on the basis of equation (8).

Figure 2



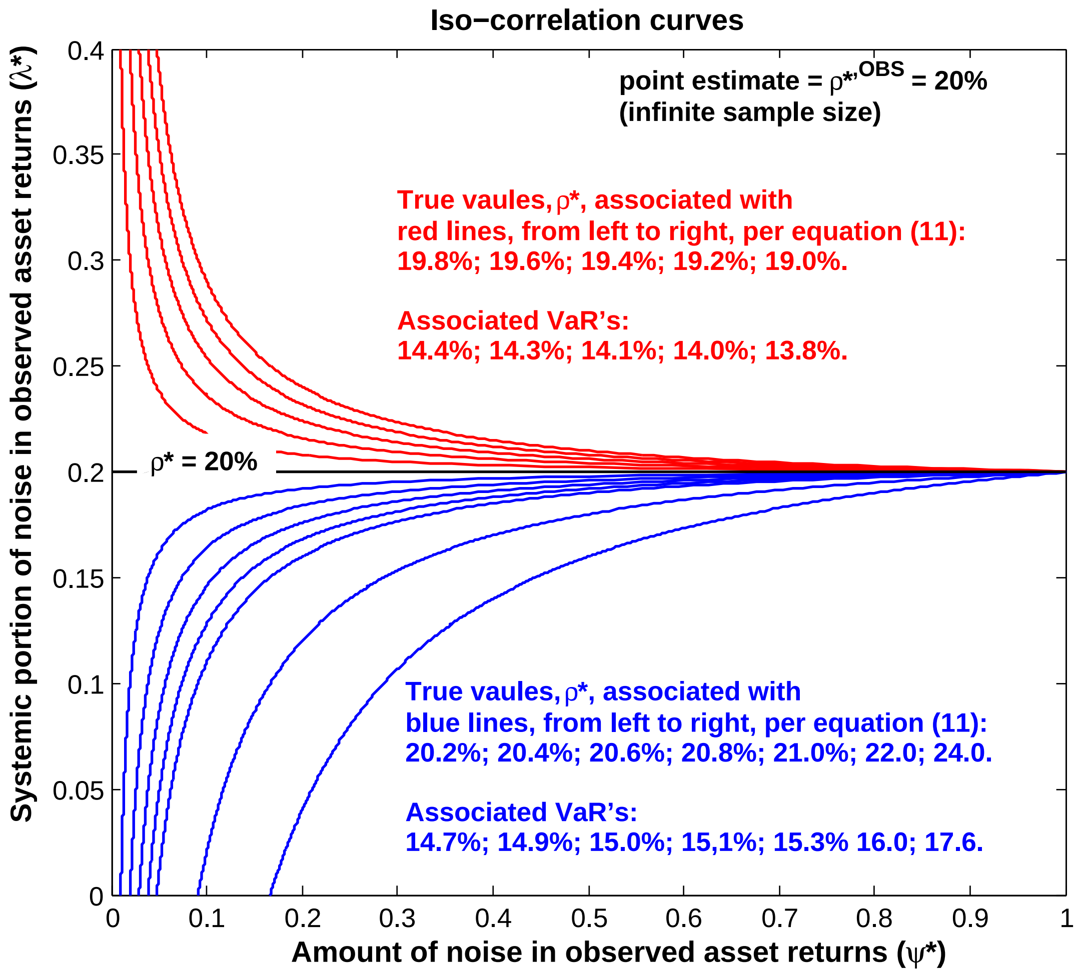
Note: All lines are plotted on the basis of equation (8).

Figure 3



Distribution of joint defaults
(an aspect of PD estimation)
typical, implied by PD* = 1%; ρ* = 20%
implied by ρ_1^c and PD(ρ_1^c)
implied by ρ_2^c and PD(ρ_1^c); ρ_2^c>ρ_1^c
implied by ρ_2^c and PD(ρ_2^c); PD(ρ_2^c)>PD(ρ_1^c)
Frequency / probability
0
0.1
0.2
0.3
0.4
0.5
0
0.005
0.01
0.015
0.02
0.025
0.03
0.035
0.04
Default rate

Figure 4


Iso−correlation curves
point estimate = ρ*'OBS = 20%
(infinite sample size)
True vaules, ρ*, associated with
red lines, from left to right, per equation (11):
19.8%; 19.6%; 19.4%; 19.2%; 19.0%.
Associated VaR's:
14.4%; 14.3%; 14.1%; 14.0%; 13.8%.
ρ* = 20%
True vaules, ρ*, associated with
blue lines, from left to right, per equation (11):
20.2%; 20.4%; 20.6%; 20.8%; 21.0%; 22.0; 24.0.
Associated VaR's:
14.7%; 14.9%; 15.0%; 15,1%; 15.3% 16.0; 17.6.
Systemic portion of noise in observed asset returns (λ*)
Amount of noise in observed asset returns (ψ*)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1