

Parameter Uncertainty and the Credit Risk of Collateralized Debt Obligations

Erik Heitfield*

Division of Research and Statistics

Federal Reserve Board

Washington, DC 20551 USA

202-452-2613

erik.heitfield@frb.gov

July 15, 2008

Abstract

This paper examines empirical challenges inherent in evaluating the credit quality of collateralized debt obligations (CDOs). In order to assess the likelihood and expected severity of a CDO tranche's losses, one needs to know the distribution of losses for each collateral asset backing the CDO, as well as the dependence of losses across collateral assets. Rating agencies and other market participants rely almost exclusively on simple copula models to describe the distribution of CDO collateral losses. For more senior CDO tranches, standard credit risk metrics such as probability of default and expected loss are highly sensitive to copula model parameters that are not directly observable. Given assumptions about the data available to a credit analyst, I compute bounds on the accuracy of copula model parameter estimators and show that in applied settings data constraints are likely to impose severe limitations on an analyst's ability to accurately evaluate CDO tranches. Thus, commonly used indicators of CDO credit quality such as credit ratings should be viewed as more preliminary and less informative than comparable indicators for unstructured corporate bonds.

Keywords: credit risk, copula, structured finance, collateralized debt obligations

JEL Classification: G24, G32, C13

*The views expressed here are those of the author and do not reflect the opinions of the Federal Reserve Board or its staff. This paper has benefited greatly from comments provided by seminar participants at the 2008 NBER Summer Institute, the Federal Reserve Board, the Hong Kong Monetary Authority, and the National University of Singapore.

1 Introduction

Collateralized debt obligations (CDOs) are structured fixed income securities whose payouts depend on the performance of pools of collateral comprised of corporate bonds or loans or other structured securities which are themselves backed by underlying collateral pools. Recently, large numbers of highly-rated CDO notes backed by mortgage securities have experienced dramatic credit rating downgrades and a few have declared defaults. As investors have become increasingly concerned about the credit risk embedded in complex structured products, demand for CDOs and other structured products has plummeted, creating liquidity and credit challenges for some financial institutions. Rating agencies that make a business of assessing the credit quality of fixed income securities have been criticized for issuing ratings for CDOs and other structured products that were too optimistic or did not capture the full range of risks associated with these securities. Perhaps most notably, the President's Working Group on Financial Markets (2008), a committee of financial regulators chaired by Treasury Secretary Henry Paulson, has concluded that “[c]redit rating agencies contributed significantly to the recent market turmoil by underestimating the credit risk of subprime residential-mortgage-backed securities and other structured credit products, notably [asset-backed] CDOs.”

Going forward, regulators and some policymakers have called upon rating agencies to treat structured credit products differently from other types of fixed income securities such as corporate bonds. For example, the President's Working Group has asked rating agencies to “make changes to the credit rating process that could clearly differentiate ratings for structured products from ratings for corporate and municipal securities” and “work with investors to provide the information investors need to make informed decisions about risk, including measures of the uncertainty associated with ratings and of potential ratings volatility.” Implicit in these recommendations is the assumption that credit ratings may not be directly comparable across different types of debt securities, and in particular, that ratings for structured securities may not be as informative as those for corporate or municipal bonds.

Why might the credit risk of similarly-rated structured and unstructured credit products differ? First, one-dimensional credit ratings convey only limited information about a security's credit risk. Typically, a bond's rating is an assessment of its likelihood of default or its expected loss relative to contractual principal and interest obligations. Other important dimensions of risk, such as the security's sensitivity to market-wide shocks, are not explicitly factored into its credit rating. Structured securities backed by diversified pools of assets are likely to be more sensitive to systematic factors than unstructured securities with similar

default probabilities or expected losses. Second, even if one focuses narrowly on those dimensions of risk captured by ratings, it may simply be more difficult for analysts to accurately gauge the credit quality of complex structured credit products. For this reason, one may need to exercise greater caution when relying on credit ratings or other metrics of structured products' future credit performance.

This paper examines the empirical challenges involved in measuring the credit risk of CDOs and shows that CDO notes are inherently more difficult to evaluate than the collateral backing them. CDO notes with significant credit enhancement only bear losses when collateral losses are substantially higher than expected, so the distribution of CDO note payouts is highly sensitive to features of the upper tail of the distribution of collateral losses. Since extreme collateral loss events are rare, accurately characterizing the upper tail of this distribution is exceedingly challenging, especially when the collateral pool consists of high quality bonds or loans that rarely default. In standard parametric portfolio credit models used by rating agencies and others to evaluate CDOs, a CDO note's loss distribution depends on measures of the likelihood and expected severity of losses for each collateral asset as well as parameters that describe the dependence of losses across collateral assets. All these parameters must be inferred from limited historical data, and small errors in estimating them can result in large errors in model-implied measures of CDO credit quality. Thus, even if we assume that the parametric model of collateral losses used to rate CDO deals is correctly specified, sampling errors in estimated model parameters are likely to translate into large errors in CDO risk metrics.

The paper is organized as follows. Section 2 briefly surveys prior research on normal copula models commonly used to assess the credit risk of CDOs and other structured credit products. Section 3 presents a stripped down version of a normal copula model and shows how it can be used to compute loss distributions for unstructured and structured securities. This model is considerably simpler than those typically used in practice. It is intended to capture the salient features of richer copula-based credit models while allowing one to investigate the role of a small number of key model parameters. Section 4 examines how changes in model parameters affect standard metrics of the credit quality of debt securities such as probability of default and expected loss. Simulations show that risk metrics for senior CDO tranches are much more sensitive to errors in model parameters than risk metrics for junior CDO tranches or unstructured bonds. Section 5 computes bounds on the accuracy of estimated model parameters given assumptions about the historical data available to an analyst. This section highlights how both the quantity and character of available data imposes quantifiable limits on an analyst's ability to accurately estimate model parameters.

Section 6 uses results from Sections 4 and 5 to simulate distributions of estimated credit risk metrics for different types of structured and unstructured bonds. These results suggest that even when high quality historical data are relatively plentiful, it may be difficult to accurately judge the credit quality of more senior CDO notes, particularly when those notes are backed by high quality collateral.

2 Literature review

Modeling dependence in realized defaults among groups of credit exposures is critical for portfolio credit risk management. Normal copula models describe default dependence using systems of correlated normal latent credit factors. Copula-based models have become popular over the last decade, both because they are computationally tractable and because they can be derived from Merton's (1974) structural debt valuation framework. Today, normal copula models are ubiquitous in risk management applications. Widely used portfolio evaluation tools developed by The RiskMetrics Group (Gupton, Finger, and Bhatia 1997) and Moody's/KMV (Kealhofer 1998) can be interpreted as variants of normal copula models (Li 2000). Normal copula models are used to compute bank regulatory requirements under the Basel II risk-based regulatory capital accord (Basel Committee on Banking Supervision 2004). Moody's, Standard and Poor's, and Fitch, the three largest bond rating agencies, rely on normal copula models to develop credit ratings for CDOs¹ and normal copula models are commonly used to price CDO notes (Andersen and Sidenius 2005a).

Despite their popularity, limitations of normal copula models have been well documented. A particular concern is that models capable of fitting observable data under typical credit conditions appear to understate the likelihood of extreme portfolio loss events. Numerous authors have proposed extensions or generalizations of the normal copula framework to address this problem. For example, Frey, McNeil, and Nyfeler (2001) propose a copula model based on thicker tailed t-distributed latent variables and Andersen and Sidenius (2005b) and Burtschell, Gregory, and Laurent (2007) extend the normal copula model to allow for unobserved heterogeneity in latent factor correlations across credit exposures.

Normal copula models and their various extensions depend on vectors of parameters that describe the probability and likely severity of individual credit loss events and the correlation structure of latent credit factors. In applied settings these parameters must be estimated from historical data or specified judgmentally by a credit analyst. Any errors in estimating

¹See Jolivet, Lassalvy, Mueller-Heumann, and Sieler (2006), Gilkes (2004), and Koo, Cromartie, and Vedova (2006) for descriptions of the normal-copula models used by Moody's, S&P, and Fitch respectively.

model parameters will naturally result in miss-measurement of the credit risk associated with individual exposures or portfolios of exposures. Simulation studies by Loffler (2001), Tarashev and Zhu (2008), and Hamerle and Rosch (2005) investigate the sensitivity of portfolio loss measures to errors in estimating copula model parameters and Fender, Tarashev, and Zhu (2008) examine how changing normal copula model parameter assumptions can affect CDO note credit ratings.

A small body of research has examined the statistical properties of normal copula model parameter estimators. Hansen and Schuermann (2006) compute confidence intervals for bond default probability parameters estimated from historical default rate data. Gordy and Heitfield (2002) compare the small sample properties of maximum likelihood and moment-based estimators of default probability and factor correlation parameters and show how imposing intuitive parameter restrictions can improve estimation efficiency. Frey and McNeil (2003) propose maximum likelihood estimators for copula model parameters in the presence of non-normal latent credit factors, and Hamerle and Rosch (2005) examine the sensitivity of model parameters to misspecification of the distribution of latent credit factors. Dullmann, Kull, and Kunisch (2008) use simulations to compare the accuracy of factor correlation estimators based on historical default data, equity returns, and market-implied asset returns.

The analysis presented in this paper contributes to the academic literature on copula models in a number of ways. I show how bounds on the accuracy of model parameters can be determined with minimal assumptions about the actual estimators used. Unlike previous research, I allow for the possibility that default probability and correlation parameters are jointly estimated either from historical default rate data alone or from a combination of default data and information on the performance of non-defaulted firms. Though others have investigated the effects of copula model specification errors on measures of portfolio credit risk, to my knowledge this paper is the first to rigorously examine how data limitations affect the accuracy of credit risk metrics for structured credit products.

3 The normal copula/beta model for correlated bond credit losses

Throughout this paper, I will use the term “simple” bonds to describe traditional debt claims on corporations or sovereigns. I will use the term “structured” bonds to describe bonds issued by CDOs, whose payouts depend on an underlying collateral pool of simple bonds. This section presents a stripped down version of the normal copula/beta model commonly

used to evaluate the credit risk of both simple and structured bonds. The specification used in this analysis allows for cross sectional correlation in realized default rates for simple bonds and stochastic losses for those bonds that default. Correlations in defaults are driven by a single systematic risk factor. For simplicity, I assume that loss rates given default are independent of the systematic factor. It should be noted that both the single systematic risk factor assumption and the assumption that there is no systematic component of loss given default can be easily relaxed, and indeed, this is commonly done in more applied settings.

3.1 Simple bonds

Under the simplest normal copula framework, bond i defaults during a specified horizon if an unobservable normal latent factor Y_i lies below the default threshold $\Phi^{-1}(\pi)$ where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative density function. The parameter π describes the bond's marginal probability of default. Cross-sectional correlation in defaults across pairs of simple bonds i and j arises from correlations in latent credit factors Y_i and Y_j . Let

$$Y_i = \sqrt{\rho}X + \sqrt{1 - \rho}E_i$$

where X is a standard normal random factor shared by all bonds, and E_i is a standard normal idiosyncratic factor that is unique for each bond. The parameter ρ lies between zero and one and determines the correlation in credit factors between pairs of bonds. Higher values of ρ imply higher correlation between credit factors, and, by extension, higher correlation in realized defaults.

Assume that bond i is a bullet loan that pays $1 + r_i$ at maturity if the obligor does not default, and $(1 + r_i)(1 - \lambda_i)$ if the obligor defaults. λ_i is a random variable describing the realized loss given default of the bond. For corporate bond exposures, this loss rate is often assumed to be drawn from a beta distribution which may or may not depend on the systematic factor(s) that drive asset correlations. As noted above, for simplicity this analysis assume that λ_i is independent of all other random variables. The beta distribution is a two parameter distribution with support on the unit interval that can be fully characterized by a mean parameter μ and a standard deviation parameter σ .²

The payout from a one dollar investment in bond i at the terminal date is

$$V_i = (1 + r_i) - 1 \{Y_i \leq \Phi^{-1}(\pi)\} \lambda_i(1 + r_i)$$

²In the statistics literature the beta distribution is most commonly characterized by two shape parameters α and β . I use the less common μ - σ parametrization to make the economic interpretation of model parameters more transparent. It can be shown that $\alpha = (\mu(1 - \mu) - \sigma^2) \frac{\mu}{\sigma^2}$ and $\beta = ((1 - \mu)^2 + \sigma^2) \frac{\mu}{\sigma^2} - 1$.

The right-most term is the realized contractual loss per dollar invested. Note that when λ_i is large, this loss rate may exceed 100 percent because of accrued but unpaid interest. Given N homogeneous bonds, the joint distribution of $V_1 \dots V_N$ is fully described by the parameter vector $\theta = (\pi, \rho, \mu, \sigma)$.

3.2 CDOs backed by simple bonds

The simplest types of collateralized debt obligations (CDOs) issue tranches of structured debt securities backed by pools of corporate bonds. The normal-copula/beta model of bond losses can be used to build up a model of CDO tranche credit losses. Consider a static CDO deal backed by N bonds. Investments are made at the “deal date” and proceeds are distributed to investors at the “terminal date”. The value of the collateral pool at the deal date is normalized to one and the value of the collateral pool at the terminal date is denoted V_p . Share c_e of the collateral pool is funded by equity investors at the deal date. The remaining $1 - c_e$ of the pool is funded by a continuum of arbitrarily thin debt tranches. Debt tranches are indexed by $c \in [c_e, 1]$. c is a tranche’s attachment point in the CDO capital structure, so higher values of c imply greater seniority. The interest paid to each debt tranche is described by the non-increasing function $r(c)$. At the terminal date, collateral is liquidated and tranche c investors are paid $1 + r(c)$ if sufficient funds are available. If V_p is not sufficient to pay all debt investors, tranches are paid according to seniority. If V_p exceeds that needed to pay debt investors, equity investors receive any residual value.

Assuming no credit losses, the total value of all debt tranches senior to tranche c is

$$\bar{V}(c) = (1 - c) + R(c)$$

where $R(c) = \int_c^1 r(s) ds$ is the total interest owed to these tranches. The realized value of tranches senior to c is

$$V(c) = \bar{V}(c) - 1 \{V_p \leq \bar{V}(c)\} (\bar{V}(c) - V_p)$$

The second right-hand term is the value of any realized credit losses for tranches senior to c . Note that the value for a “slice” of the CDO with attachment point c_l and detachment point c_h is $V(c_l) - V(c_h)$. The value of the equity tranche is $V_p - V(c_e)$.

To keep notation simple, this analysis is restricted to CDOs backed by equal-weighted pools of bonds that are homogeneous in the sense that all bonds in the pool share the same parameter vector θ and pay the same interest rate r_p . Let $M = \sum_{n=1}^N 1 \{Y_n \leq \Phi^{-1}(\pi)\}$ be a

random variable that describes the number of bonds in the CDO collateral pool that default by the terminal date, and let $\bar{\lambda}_M = \frac{1}{M} \sum_{j=1}^M \lambda_j$ be the average loss given default for those M bonds. The value of the collateral pool at the terminal date is

$$V_p = (1 + r_p) - \frac{M}{N} \bar{\lambda}_M (1 + r_p).$$

The random variables M and $\bar{\lambda}_M$ determine V_p . M is a draw from a binomial-normal mixture distribution, and, conditional on M , $\bar{\lambda}_M$ is an average of M independent beta random variables. Neither the marginal distributions of M nor the conditional distribution of $\bar{\lambda}_M$ given M can be expressed in closed form, but both can be computed analytically with high precision. The product of these two distributions is the joint distribution of M and $\bar{\lambda}_M$, which provides all the information necessary to compute the joint distribution of V_p and $V(c)$ for all c .

The distribution of CDO tranche payouts is fully determined by N , r_p , $r(c)$, and the normal copula/beta model parameter vector θ for the collateral pool. N , r_p , and $r(c)$ are known features of the CDO contract, but θ is not directly observable by market participants. Given θ , any number of relevant metrics of the credit risk associated with a CDO note can be computed. The next section examines how three common metrics of credit risk depend on θ .

4 Credit risk metric sensitivity to model parameters

This analysis considers three standard metrics of credit quality: probability of default, expected loss, and conditional expected loss. Define the expectation operator $E[Z]$ as the expected value of the random variable Z whose distribution is determined by θ . Let V be the value of a one dollar investment in a debt security at the terminal date and let r be the contractual interest on that security. The security's probability of default is defined as

$$PD = E[1 \{V < 1 + r\}]. \tag{1}$$

PD describes the likelihood of a credit loss, but not the magnitude of the loss. The expected loss

$$EL = (1 + r) - E[V] \tag{2}$$

summarizes the expected likelihood and the magnitude of a credit loss. Note that EL may exceed 100 percent because both principal and accrued interest may be lost.

PD and EL describe the first moment of a security's loss distribution. In portfolio risk management applications such as economic capital allocation, analysts also require information about a security's marginal contribution to portfolio-wide losses. A number of risk metrics useful for describing the dependence between an individual exposure's credit losses and those of a broader portfolio have been proposed in the risk management literature, and I do not propose to survey them here. This analysis will consider one such measure derived from an asymptotic single risk factor approximation. Gordy (2003) shows that if a portfolio is well diversified and its overall loss rate depends on a single systematic factor \bar{X} then an exposure's marginal contribution to portfolio value-at-risk (VaR) can be determined analytically by calculating the conditional expected loss of the exposure given an adverse draw of the systematic risk factor. The conditional expected loss associated with a q th percentile portfolio VaR measure is

$$EL_q = (1 + r) - E[V | \bar{X} = \bar{x}_q] \quad (3)$$

where \bar{x}_q is the $1 - q$ th percentile of the stochastic systematic risk factor. Unlike PD and EL , which describe the center of the distribution of losses, EL_q describes the tail of this distribution.

Under the normal copula/beta model, EL , PD , and EL_q for both simple and structured bonds are determined by the parameter vector θ . For simple bonds,

$$PD = \pi$$

and

$$EL = (1 + r)\pi\mu.$$

If we assume that correlation between the single systematic factor underlying a financial institution's overall asset portfolio \bar{X} and the systematic factor that affects bond default rates X is 50 percent, the conditional expected loss for a simple bond is

$$EL_q = (1 + r)\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{0.5\rho}\bar{x}_q}{\sqrt{1 - 0.5\rho}}\right)\mu$$

where $\bar{x}_q = \Phi^{-1}(1 - q)$.³

For simple bonds, PD is determined by the normal copula marginal default probability parameter π , EL depends on both π and the expected loss-given-default parameter μ , and EL_q is a function of π , μ , and the asset value correlation parameter ρ . For structured bonds,

³See Pykhtin and Dev (2002) for detailed analysis of conditional expected loss in the presence of a single portfolio-wide systematic factor and a correlated CDO collateral-pool-specific factor.

simple analytic formulas for PD , EL , and EL_q are not available, but these risk metrics can be computed numerically for any value of θ . In contrast to the case for simple bonds, PD , EL , and EL_q for structured bonds each depend on all four elements of θ .

The true value of the normal copula/beta model parameters, which will be denoted with the subscript “0”, cannot be directly observed by a credit analyst. In practice, model parameters are either determined judgmentally or estimated from historical data. Any differences between θ_0 and the value of θ used to compute the expectations in equations (1), (2), and (3) can result in errors in imputed risk metrics.

To illustrate how errors in θ can affect imputed risk metrics, this paper examines two hypothetical CDO deals summarized in Tables 1 and 2. Both CDO deals are backed by homogeneous collateral pools of 100 simple bonds described in the bottom rows of the tables. In the mezzanine CDO example $\pi_0 = 0.05$, $\rho_0 = 0.20$, $\mu_0 = 0.55$, and $\sigma_0 = 0.35$ and bonds have a maturity of 5 years. In the high-grade CDO example, all normal copula/beta parameters are the same except the default probability parameter, which is set at $\pi_0 = 0.01$. The default probability parameter for the mezzanine and high-grade cases roughly corresponds to the five-year cumulative default rates for corporate bonds rated Baa- and A- respectively. The collateral’s asset value correlation and LGD parameters are chosen to be broadly consistent with those used in major rating agencies’ or regulators’ credit risk models.⁴

In both cases, the collateral pool is financed by five debt tranches and an equity tranche. As is typical of securitization deals, the credit risk metrics for a particular debt tranche bear little direct relation to those of the underlying collateral pool, but are very sensitive to the tranche’s position in the CDO capital structure. Tranches with lower attachment points, which are the first to take losses, have much higher default probabilities, expected losses, and conditional expected losses than more senior tranches. In the two examples presented here, tranche attachment and detachment points are chosen so that tranches in each deal with the same seniority have similar risk metrics. Because the high-grade deal is backed by safer collateral, each tranche of that deal requires less credit support to achieve a given default probability or expected loss. A constant risk-free interest rate of four percent is assumed, and spreads for CDO collateral and all tranches are set so that a risk-neutral pricing model based on θ_0 would value all bonds at par at the deal date.

Figures 1 through 4 show how deviations in each component of θ from θ_0 affects PD ,

⁴Basel II uses corporate bond asset correlation parameters ranging from 0.12 to 0.24 depending on firm size (Basel Committee on Banking Supervision 2004). The LGD parameters are taken from default values for unsecured corporate bonds used in rating agency CDO models.

Table 1: Credit risk statistics for hypothetical CDO deal backed by 100 mezzanine-rated bonds.

Tranche	Position (%)	Spread (bp)	PD (%)	EL (%)	$EL_{0.95}$ (%)
Junior	3.5 – 6.5	188	14.06	9.86	39.85
Jr. Mezz.	6.5 – 9.0	77	5.36	3.92	18.11
Sr. Mezz.	9.0 – 12.0	34	2.54	1.73	8.04
Senior	12.0 – 15.0	14	1.07	0.71	3.07
Sup. Sen.	15.0 – 100.0	0	0.43	0.01	0.05
Collateral		56	5.00	3.45	8.14

Table 2: Credit risk statistics for hypothetical CDO deal backed by 100 high-grade bonds.

Tranche	Position (%)	Spread (bp)	PD (%)	EL (%)	$EL_{0.95}$ (%)
Junior	1.0 – 2.0	183	13.24	9.56	35.17
Jr. Mezz.	2.0 – 3.0	76	5.41	3.89	17.06
Sr. Mezz.	3.0 – 4.0	37	2.53	1.87	8.62
Senior	4.0 – 6.0	15	1.29	0.75	3.35
Sup. Sen.	6.0 – 100.0	0	0.39	0.01	0.03
Collateral		11	1.00	0.68	1.92

Figure 1: Sensitivity of mezzanine CDO risk metrics to errors in collateral default parameters π and ρ .

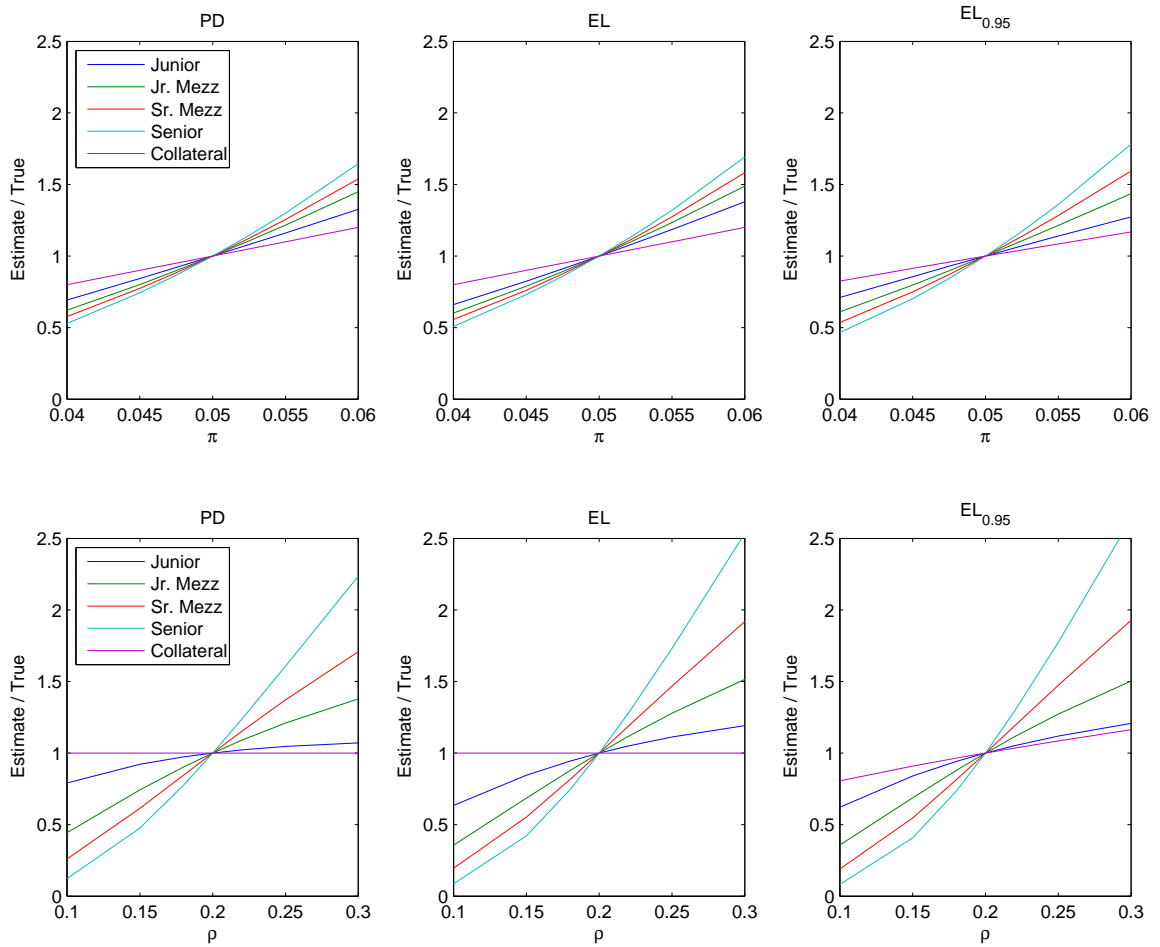


Figure 2: Sensitivity of mezzanine CDO risk metrics to errors in loss severity parameters μ and σ .

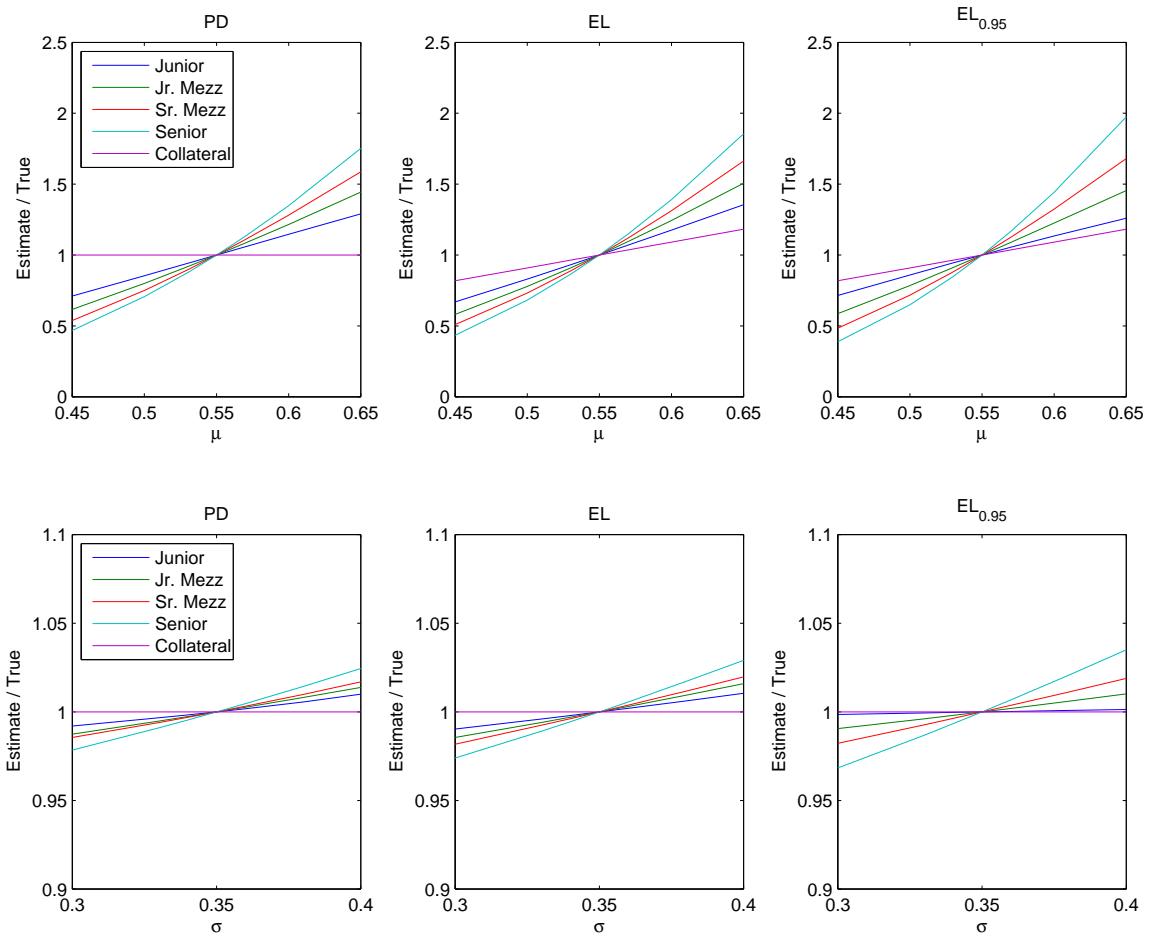


Figure 3: Sensitivity of high-grade CDO risk metrics to errors in collateral default parameters π and ρ .

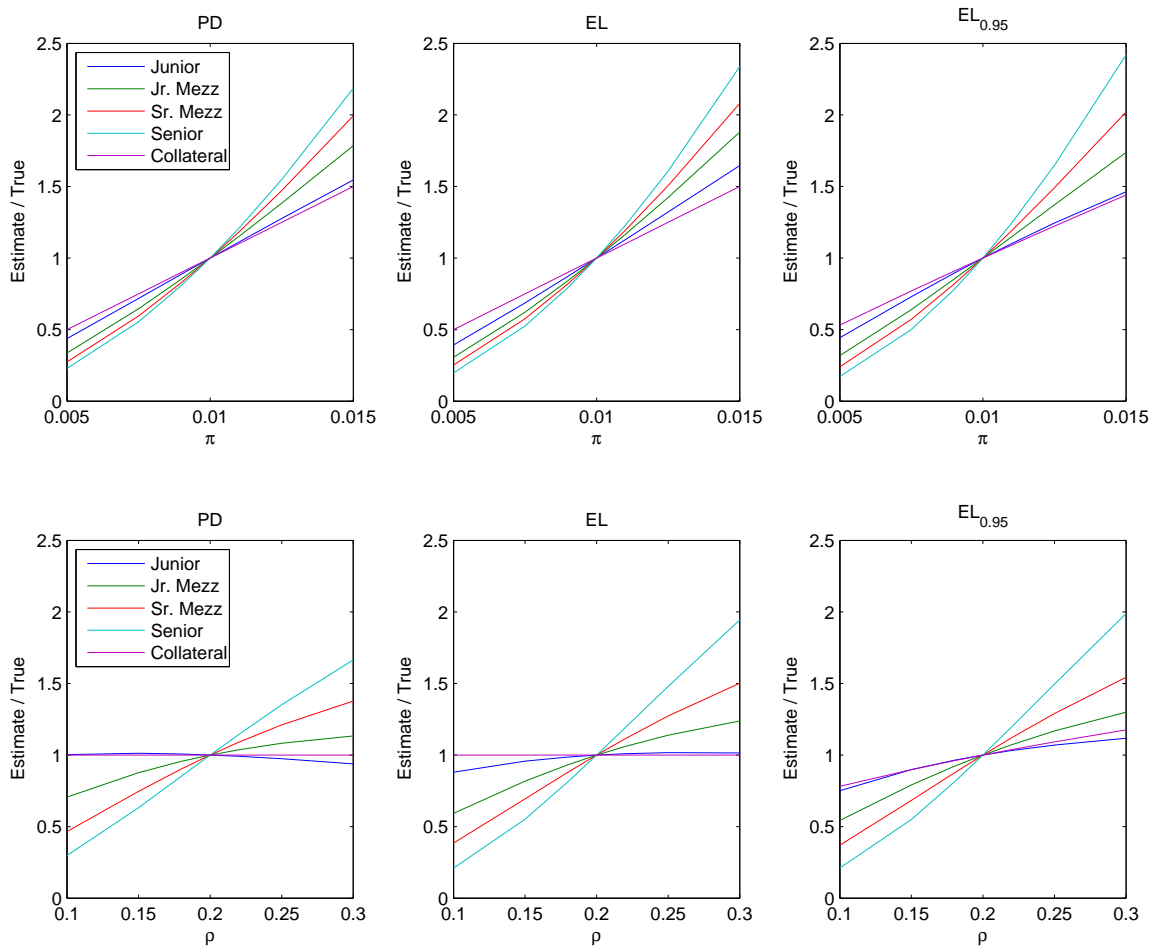
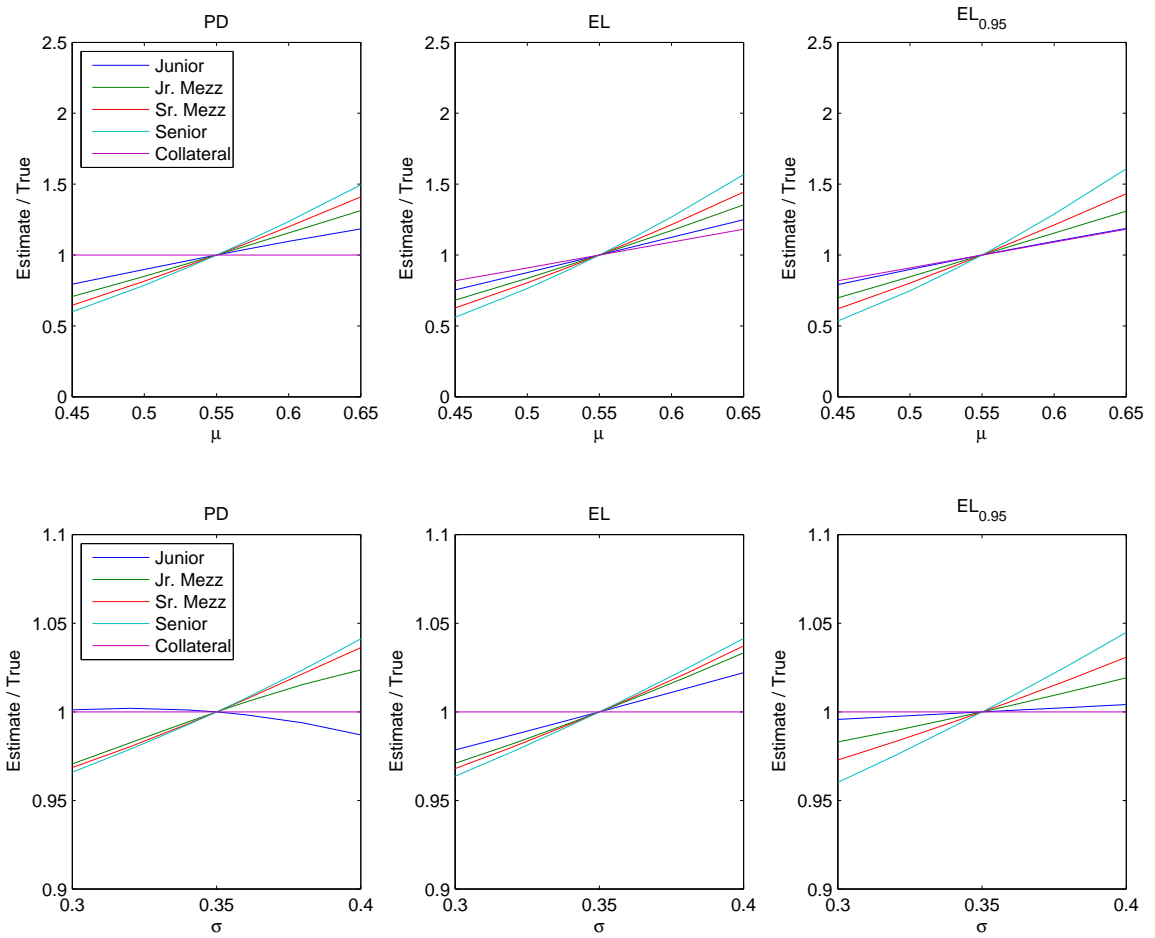


Figure 4: Sensitivity of high-grade CDO risk metrics to errors in loss severity parameters μ and σ .



EL , and EL_q . Each line in a panel plots the ratio of a particular risk metric to its true value (i.e., the risk metric computed given $\theta = \theta_0$) as θ changes. For example, the first panel of Figure 1 shows how bonds' implied default probabilities change with π , holding all other elements of θ fixed at their true values. All lines cross at the true value of the parameter in question.

As these results plainly show, structured bonds are considerably more sensitive to errors in each component of θ than simple bonds, and the higher is a bond's position in the CDO capital structure the greater is its sensitivity to parameter errors. For more senior tranches, risk metrics are remarkably sensitive to model parameters. For example, if a credit risk manager were to use an asset value correlation parameter of 15 percent to evaluate the "Senior" tranche of the mezzanine CDO in an environment where the true correlation was 20 percent, she would understate the tranche's risk metrics by over 50 percent. Conversely, if she used an asset correlation parameter of 25 percent, she would overstate the tranche's risk metrics by 75 percent.

5 Bounding the accuracy of model parameters

The sensitivity of structured bond risk metrics to θ begs the question, how accurately can model parameters be estimated? The answer, of course, depends on the type and volume of historical data available and the statistical estimator used. Practitioners typically use historical data on the default frequencies of rated bonds to estimate default probability parameters. Factor correlation parameters can also be estimated from default data by backing out implied factor correlation parameters from cross-sectional correlations in observed default rates.⁵ However, as Gordy and Heitfield (2002) show, accurately estimating correlation parameters from default rate data requires long data histories and/or strong *ex ante* parameter restrictions. As noted earlier, the latent credit factors of the normal copula model have a structural interpretation as obligor asset returns in a Merton (1974) valuation framework. Leveraging this fact, practitioners often use information on equity returns or imputed asset returns for publicly traded firms as proxies for latent credit factors.⁶ Alternatively, some practitioners use information on bond rating transitions (which are far more common than

⁵The corporate bond factor correlation parameters used in version 3.0 of Standard and Poor's CDO Evaluator copula-based rating model are estimated using this approach (Gilkes 2004, pg. 8-9).

⁶For example, the corporate bond factor correlation parameters used in version 3.0 of Fitch Rating's Vector copula-based CDO model are estimated using a factor model of equity returns included in the Dow Jones Global Universe index (Koo, Cromartie, and Vedova 2006, pg. 21).

bond defaults) to estimate factor correlations.⁷ Loss given default parameters are typically estimated from data on the market price of traded debt shortly after a default event or ultimate recoveries on defaulted bonds or loans.

In this section, I present a stylized data generating process (DGP) capable of describing the range of data that might reasonably be available to a credit analyst. I assume that the normal copula/beta model is correctly specified in the sense that the DGP is consistent with this model given a true parameter vector θ_0 . Using the DGP, I derive lower bounds on the sampling variance of any unbiased estimator of θ_0 . This allows me to investigate how the quantity and character of available historical data affects the accuracy of normal copula parameter estimates.

Assume that an analyst observes T cohorts that each contain N simple bonds. All bonds in all cohorts share the same underlying model parameter vector θ_0 . Within a cohort, all bonds are sensitive to the same systematic risk factor X , but systematic risk factors are assumed to be independent across cohorts. Such data could arise, for example, if one observed information on T non-overlapping cohorts of bonds over time. In applied settings, the stylized assumption that historical data can be grouped into independent, homogeneous cohorts could be easily generalized in useful ways. For example, a more realistic DGP might allow for correlations across cohort-specific systematic factors since in practice cohorts may well be overlapping in time. Within cohort heterogeneity, particularly with respect to π_0 , could also be accommodated as in Gordy and Heitfield (2002). This would allow one to investigate the costs and benefits of pooling historical data across different types of credit exposures. For the purpose at hand, relatively stylized DGP assumptions are more useful because they allow us to investigate how broad features of available data affect parameter estimates.

For each bond i in cohort t , the analyst observes the following information *ex post*. First the analyst observes an indicator variable D_{it} which is equal to one if the bond defaults (i.e. if $Y_{it} \leq \Phi^{-1}(\pi_0)$) and zero otherwise. Second, if bond it defaults, the analyst observes the realized loss rate given default λ_{it} . Finally, if the bond does not default, the analyst observes a noisy signal of the bond's realized latent credit factor Y_{it} , denoted Y_{it}^* . This signal is meant to capture that information contained in proxy data on realized latent credit factors such as equity returns, imputed asset returns, or rating migrations.

Y_{it}^* is a weighted sum of Y_{it} and a standard normal error term U_{it} which is assumed to

⁷Moody's uses both rating transition information and correlations in imputed asset returns to calibrate latent factor correlations used in version 2.3 of its CDOROM rating model (Fu, Gluck, Mazataud, and Rosa 2004). It relies almost exclusively on historical rating transitions to calibrate correlations for asset-backed securities that serve as CDO collateral (Toutain, Rosa, Fu, Mazataud, and Jolivet 2005)

be independent of all other variables in the model:

$$Y_{it}^* = \sqrt{\omega} \left(\sqrt{1-\psi} Y_{it} + \sqrt{\psi} U_{it} \right).$$

The parameter ψ lies on the unit interval and captures the relative amount of noise in the signal Y_{it}^* . In the limiting case where $\psi = 0$, Y_{it}^* is perfectly correlated with the realized credit factor Y_{it} . At the other extreme, where $\psi = 1$, Y_{it}^* provides no information about Y_{it} . In this case Y_{it}^* can be safely ignored by the analyst, and inference about θ_0 must be based solely on information about whether or not bonds have defaulted. ω controls the scale of Y_{it}^* . The nuisance parameters ω_0 and ψ_0 are not of direct interest to the analyst, but they must be estimated in order to use information from Y_{it}^* to estimate π_0 and ρ_0 .⁸ For the remainder of this paper we will redefine θ to include these nuisance parameters.

Conditional on X , the likelihood of observing the vector $(D_{it}, Y_{it}^*, \lambda_{it})$ can be expressed as the product of a marginal likelihood for (D_{it}, Y_{it}^*) given X which depends on π , ρ , ψ , and ω and a conditional likelihood for λ_{it} given D_{it} which depends on μ and σ :

$$f(D_{it}, Y_{it}^*, \lambda_{it} | X; \theta) = f(D_{it}, Y_{it}^* | X; \pi, \rho, \psi, \omega) f(\lambda_{it} | D_{it}; \mu, \sigma) \quad (4)$$

The first right-hand term in (4) is

$$\begin{aligned} f(D_{it}, Y_{it}^* | X; \pi, \rho, \psi, \omega) &= \left[\Phi \left(\frac{\Phi^{-1}(\pi) - \sqrt{\rho} X}{\sqrt{1-\rho}} \right) \right]^{D_{it}} \\ &\times \left[\phi \left(\frac{(Y_{it}^*/\sqrt{\omega}) - \sqrt{1-\psi} \sqrt{\rho} X}{\sqrt{S}} \right) \right]^{1-D_{it}} \\ &\times \left[\Phi \left(-\frac{S \Phi^{-1}(\pi) - \psi \sqrt{\rho} X - (1-\rho) \sqrt{1-\psi} (Y_{it}^*/\sqrt{\omega})}{\sqrt{S} \sqrt{\psi(1-\rho)}} \right) \right]^{1-D_{it}} \end{aligned} \quad (5)$$

where $S = (1-\psi)(1-\rho) + \psi$. See the appendix for a derivation of (5). The second term in (4) is the likelihood of the loss rate λ_{it} , which is only observable if $D_{it} = 1$. It can be written

$$f(\lambda_{it} | D_{it}; \mu, \sigma) = [\beta(\lambda_{it})]^{D_{it}}$$

where $\beta(z)$ is the density function for a beta-distributed random variable with mean and

⁸Exceptions occur in the limiting cases where $\psi_0 = 0$ or $\psi_0 = 1$. When $\psi_0 = 0$ there is no uncertainty about the information content of Y_{it}^* and hence no need to estimate ψ_0 . In this case it is still necessary to estimate the scale parameter ω_0 . When $\psi_0 = 1$, Y_{it}^* is irrelevant to inference of θ , so there is no need to estimate either ψ_0 or ω_0 .

standard deviation parameter μ and σ respectively. The joint likelihood for cohort t is

$$L_t(\theta) = \int \prod_{i=1}^N f(D_{it}, Y_{it}^* | x; \pi, \rho) \phi(x) dx \prod_{i=1}^N f(\lambda_{it} | D_{it}; \mu, \sigma) \quad (6)$$

Given the likelihood of the DGP, the well known Cramer-Rao lower bound defines the minimum covariance matrix of any asymptotically unbiased estimator for θ_0 . The Fischer information matrix for our DGP is

$$I(\theta_0) = \text{E} \left[\frac{\partial^2 \ln L_t(\theta_0)}{(\partial \theta)^2} \right].$$

If an estimator of θ_0 is asymptotically unbiased, then the difference between its covariance matrix and $\frac{1}{T} (I(\theta_0))^{-1}$ is a positive semi-definite matrix. Given θ_0 , the Cramer-Rao bound can be computed directly for any combinations of T and N .⁹ Figures 5 and 6 show how the minimum standard deviation of π , ρ , μ , and σ vary as N , T , and ψ_0 change. Since the sampling errors of μ and σ do not depend on ψ_0 , the bottom two panels of the figures only show how estimator standard deviations change with N and T .¹⁰

Though the absolute magnitude of the sampling error of π is larger for the mezzanine bond population than for the high-grade population, the size of this error relative to the value of π_0 is larger for the high-grade population. Simply put, it is more difficult to accurately estimate the frequency of low probability events. Somewhat surprisingly, the sampling error of π can be significantly reduced by incorporating available information on latent credit factors. Given that default probability parameters are commonly calibrated using only long-run default frequencies, this fact does not appear to be widely appreciated.

When latent credit factors are not fully observable, the sampling error of ρ is inversely related to the population default probability parameter π_0 . The standard deviations for ρ in the high-grade bond population with unobservable latent factors ($\psi_0 = 1.0$) are roughly double those for the mezzanine bond population. However, when information on latent credit factors is available, the relationship between π_0 and the sampling error of ρ is more attenuated. In the limit, where latent credit factors for non-defaulted bonds are fully ob-

⁹For a given data vector the hessian matrix is computed from (6). The hessian's expectation is approximated by computing the average of the Hessians produced by 10,000 simulated data vectors drawn from the assumed DGP.

¹⁰Because the marginal distribution of D_{it} does not depend on μ and σ and the conditional distribution of λ_{it} does not depend on π , ρ , ψ , and ω , the information matrix is block diagonal with respect to the two sub-vectors. If an estimator achieves the Cramer-Rao bound, the sub-vectors will be uncorrelated with one another. Note that this result does not hold in the more general setting in which loss given default depends on X .

Figure 5: Minimum standard deviation of unbiased normal copula parameter estimators for mezzanine bonds under various data generating processes ($\pi_0 = 0.05$, $\rho_0 = 0.20$, $\mu_0 = 0.55$, $\sigma_0 = 0.35$).

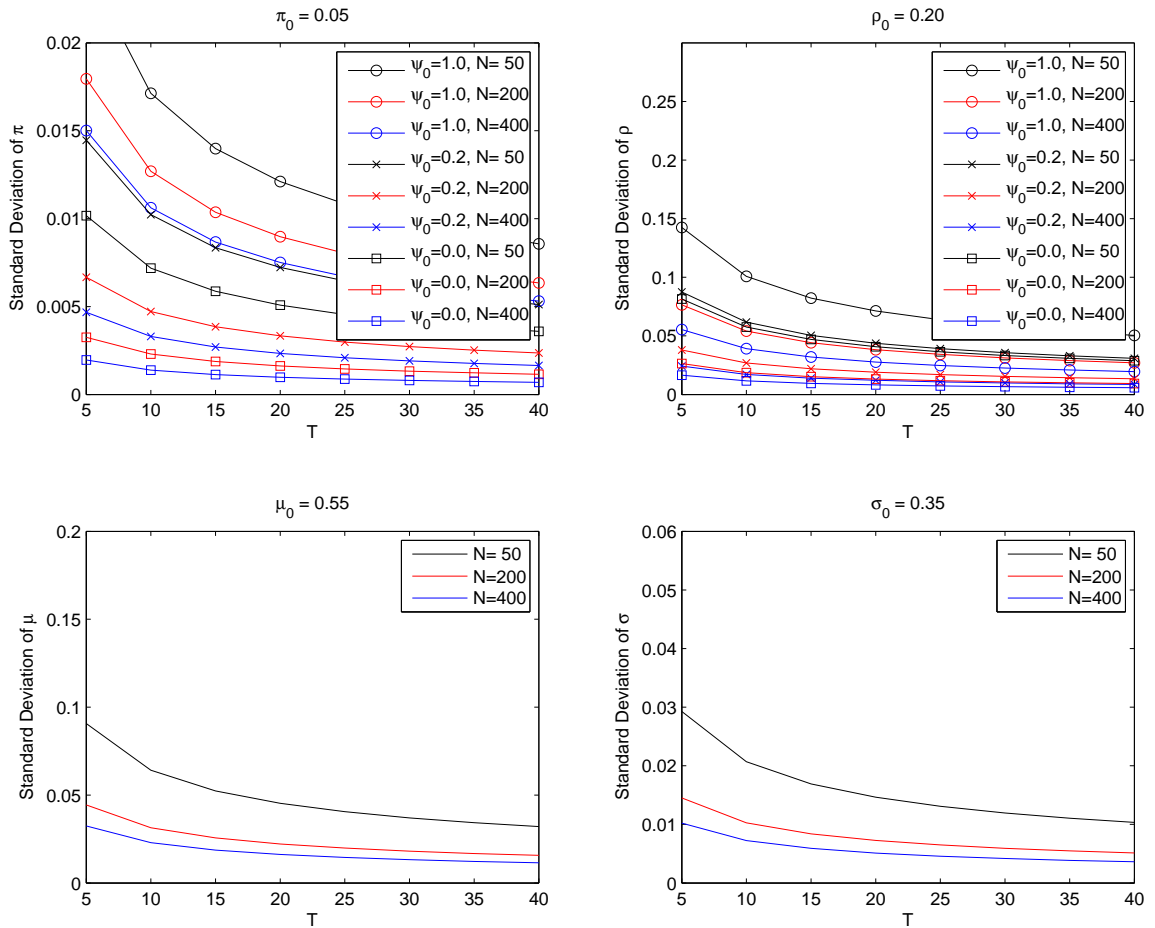
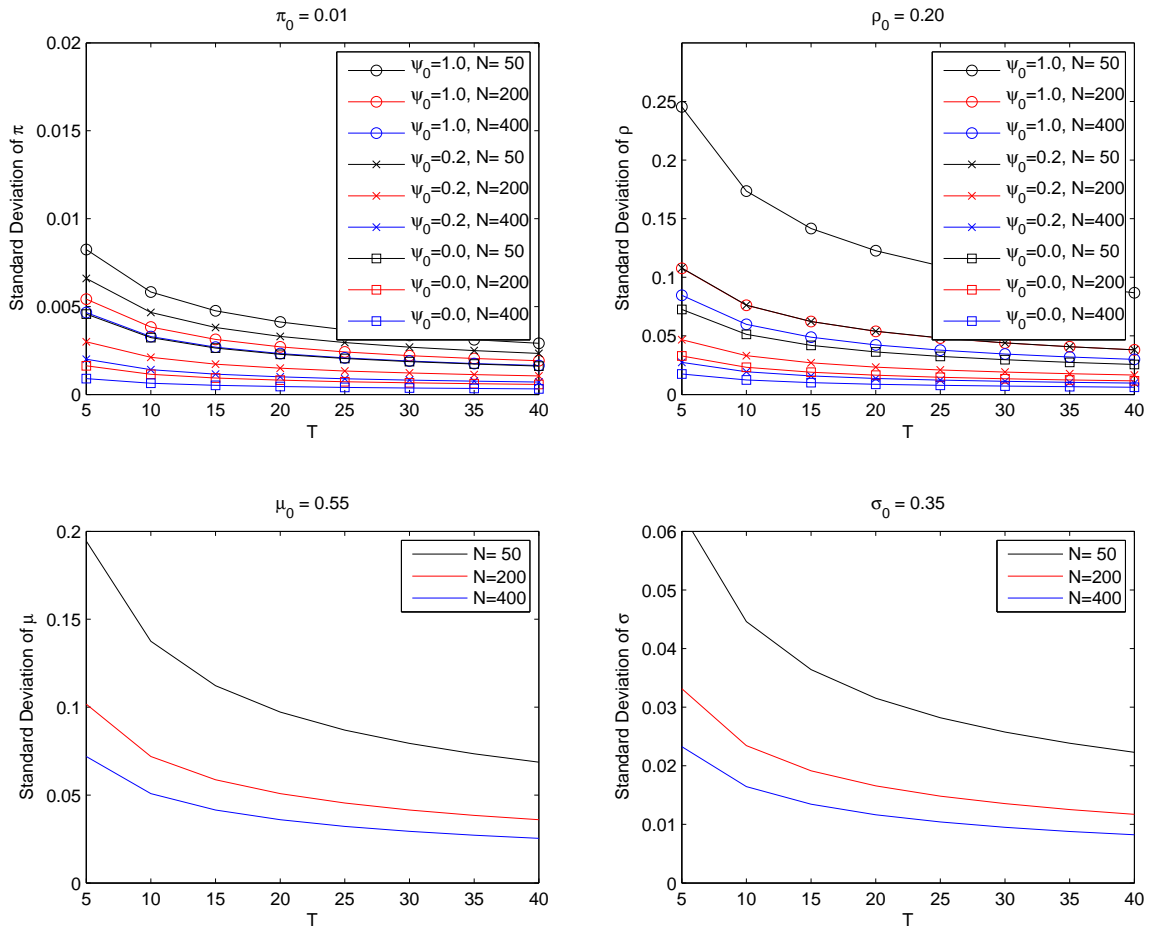


Figure 6: Minimum standard deviation of unbiased normal copula parameter estimators for high-grade bonds under various data generating processes ($\pi_0 = 0.01$, $\rho_0 = 0.20$, $\mu_0 = 0.55$, $\sigma_0 = 0.35$).



servable ($\psi_0 = 0.0$), the standard deviations of ρ are similar for high-grade and mezzanine populations. These results underscore the value of observing high quality historical data on the performance of those bonds that do not default. Comparing results for $\psi_0 = 1.0$ and $\psi_0 = 0.2$ we see that when estimating ρ , observing even somewhat noisy information on latent credit factors is comparable to a four-fold increase in cohort size. Indeed, for the high-grade bond population, estimating ρ with reasonable precision virtually requires some type of credit factor data.

As one would expect, estimates of μ and σ are more accurate for populations where high default rates imply more plentiful historical data on default loss severities. Because loss severities are assumed to be independent of the systematic risk factor, increasing the cohort size N has exactly the same effect on the standard errors of μ and σ as increasing the number of cohorts T .

6 Distribution of credit risk metrics

Section 4 shows how standard measures of credit risk depend on estimated normal copula model parameters, and Section 5 shows how the accuracy of parameter estimates depends on the data available to an analyst. This section combines these results to examine how the distributions of PD , EL and EL_q are affected by characteristics of the data generating process.

Given the sampling distribution of an estimator θ of θ_0 , the distribution of PD , EL , and EL_q for simple and structured bonds can be estimated using Monte Carlo simulation. Unfortunately, while the Fischer information inequality allows one to bound the sampling variance of unbiased estimators, it provides no additional insights into the small sample properties of such estimators. Hence, in order to simulate the distribution of PD , EL , and EL_q , we need to make some assumptions about the sampling distribution of θ . The parameters π , ρ , and μ , are bounded between zero and one and σ is bounded between zero and $\sqrt{(1-\mu)\mu}$, so the natural assumption that θ is drawn from a multivariate normal distribution is inappropriate. One can circumvent this problem by reparameterizing the normal copula model in such a way that each model parameter is defined over the entire real line. Let

$$\tilde{\theta} = \begin{bmatrix} \Phi^{-1}(\pi) \\ \Phi^{-1}(\rho) \\ \Phi^{-1}(\mu) \\ \Phi^{-1}\left(\frac{\sigma}{\sqrt{(1-\mu)\mu}}\right) \end{bmatrix}.$$

The transformed parameter vector $\tilde{\theta}$ has support on \mathfrak{R}^4 , and the sampling variance for a minimum variance unbiased estimator of $\tilde{\theta}_0$ can be derived by inverting the Fischer information matrix for the reparameterized likelihood function.¹¹

The sampling distributions of PD , EL , and EL_q are simulated as follows. For each Monte Carlo iteration I draw a value of $\tilde{\theta}$ from a multivariate normal distribution with mean and covariance parameters determined by $\tilde{\theta}_0$ and the Cramer-Rao bound. Using this parameter value, I compute implied PD , EL , and EL_q for simple bonds and CDO tranches. The distribution of 10,000 simulated values of PD , EL , and EL_q approximates the sampling distribution of these risk metrics given a minimum variance unbiased estimator of $\tilde{\theta}$.

Simulations are run for both the mezzanine and high-grade CDO deals described in Section 4 under four hypothetical data generating processes:

- a “wide” panel with partially observable credit factors ($N = 400$, $T = 5$, $\psi = 0.2$);
- a “long” panel with partially observable credit factors ($N = 100$, $T = 20$, $\psi = 0.2$);
- a “wide” panel with unobservable credit factors ($N = 400$, $T = 5$, $\psi = 1.0$); and
- a “long” panel with unobservable credit factors: ($N = 100$, $T = 20$, $\psi = 1.0$).

To put these DGP assumptions in perspective, note that Moody’s publishes what are probably the most comprehensive historical data on corporate bond rating performance. This dataset includes tabulations of rating transitions and defaults by cohort from 1970 to the present. Given the five-year deal horizon used in our study, the Moody’s dataset would span 7 non-overlapping cohorts. On average, each of these cohorts contains about 400 obligors per initial rating grade. Data on loss severities tend to be more limited. For example, S&P’s LossStats database contains information on bond and loan recovery rates for US corporate default events going back to 1988.

Tables 3 and 4 report the bias and root-mean-squared-error for estimated risk metrics. To make it easier to compare the relative magnitude of errors across tranches with different risk metrics, these results are presented as percentage deviations from the true values. Tables 5 and 6 report 90 percent confidence intervals for estimated risk metrics. Figures 7 through 14 plot distributions of estimated risk metrics for collateral bonds and selected CDO tranches under each DGP.

¹¹An unbiased estimator of $\tilde{\theta}_0$ is not an unbiased estimator of θ_0 . For the DGP assumptions used in this analysis, the bias in θ implied by an unbiased $\tilde{\theta}$ is less than ten percent for all parameters in all cases, and is usually less than five percent.

Table 3: Bias and root-mean-squared-error of risk metrics for a mezzanine CDO.

	<i>PD</i>		<i>EL</i>		<i>EL</i> _{0.95}	
	Bias	RMSE	Bias	RMSE	Bias	RMSE
<i>N</i> = 100, <i>T</i> = 20, $\psi_0 = 0.2$						
Junior	0.2	20.2	1.0	25.6	-0.1	21.4
Jr. Mezz	2.5	33.5	3.8	38.8	2.6	36.3
Sr. Mezz	5.6	46.0	7.9	53.2	7.7	53.8
Senior	11.2	63.6	14.6	73.1	17.0	79.8
Collateral	0.5	10.2	0.4	11.6	0.6	13.8
<i>N</i> = 400, <i>T</i> = 5, $\psi_0 = 0.2$						
Junior	0.4	17.7	0.9	21.6	-0.0	17.7
Jr. Mezz	1.8	27.2	2.6	30.9	1.7	28.9
Sr. Mezz	3.6	35.8	4.9	40.5	4.8	41.3
Senior	6.6	47.0	8.5	53.0	10.0	58.1
Collateral	0.3	8.9	0.3	10.6	0.4	11.2
<i>N</i> = 100, <i>T</i> = 20, $\psi_0 = 1.0$						
Junior	3.2	35.7	4.8	42.4	0.8	32.2
Jr. Mezz	7.6	52.9	9.9	60.3	6.5	53.4
Sr. Mezz	13.4	70.8	17.5	82.1	16.5	80.6
Senior	23.8	99.1	30.3	115.5	34.5	125.9
Collateral	2.1	21.5	2.1	22.2	1.7	21.0
<i>N</i> = 400, <i>T</i> = 5, $\psi_0 = 1.0$						
Junior	9.6	50.4	10.8	53.8	2.8	32.8
Jr. Mezz	11.9	57.6	12.4	59.2	7.1	47.1
Sr. Mezz	12.8	61.4	13.4	63.5	11.6	60.9
Senior	14.2	67.0	15.6	71.1	17.5	78.6
Collateral	3.9	31.4	3.8	31.9	1.2	20.0

All figures expressed as percentage deviations from the true values.

Table 4: Bias and root-mean-squared-error of risk metrics for a high-grade CDO.

	<i>PD</i>		<i>EL</i>		<i>EL</i> _{0.95}	
	Bias	RMSE	Bias	RMSE	Bias	RMSE
<i>N</i> = 100, <i>T</i> = 20, $\psi_0 = 0.2$						
Junior	0.0	26.1	1.7	31.7	-0.1	27.0
Jr. Mezz	3.3	39.9	5.1	46.0	3.3	42.3
Sr. Mezz	7.3	54.1	9.4	60.6	8.5	59.4
Senior	12.3	69.3	17.3	83.4	19.0	88.0
Collateral	1.8	22.2	1.5	24.8	1.7	26.3
<i>N</i> = 400, <i>T</i> = 5, $\psi_0 = 0.2$						
Junior	0.6	25.3	2.1	30.2	0.2	24.7
Jr. Mezz	3.5	37.1	4.9	42.1	3.0	38.0
Sr. Mezz	6.6	48.6	8.1	53.7	7.1	52.3
Senior	10.2	60.4	13.6	71.0	14.9	75.0
Collateral	1.8	19.8	1.5	23.6	1.5	23.5
<i>N</i> = 100, <i>T</i> = 20, $\psi_0 = 1.0$						
Junior	1.8	41.6	4.1	48.7	-0.4	38.2
Jr. Mezz	7.0	60.0	10.1	69.3	6.2	61.1
Sr. Mezz	14.6	82.3	19.1	94.1	17.1	90.1
Senior	25.5	110.5	37.9	140.5	41.0	147.7
Collateral	4.8	36.4	4.7	39.0	4.4	40.2
<i>N</i> = 400, <i>T</i> = 5, $\psi_0 = 1.0$						
Junior	9.2	58.1	11.9	64.8	1.7	38.7
Jr. Mezz	14.6	73.7	16.2	78.8	7.2	56.8
Sr. Mezz	18.0	85.4	19.4	90.1	13.6	76.2
Senior	21.3	96.6	25.0	107.6	24.6	107.4
Collateral	7.9	47.5	7.6	49.4	2.9	35.5

All figures expressed as percentage deviations from the true values.

Table 5: 90% confidence intervals for risk metrics of a mezzanine CDO.

Tranche	PD (%)			EL (%)			$EL_{0.95}$ (%)		
	True	5th	95th	True	5th	95th	True	5th	95th
$N = 100, T = 20, \psi_0 = 0.2$									
Junior	14.06	9.53	18.90	9.86	6.05	14.35	39.85	25.92	53.91
Jr. Mezz	5.36	2.82	8.68	3.92	1.88	6.84	18.11	8.74	30.36
Sr. Mezz	2.54	1.07	4.86	1.73	0.64	3.60	8.04	2.87	16.77
Senior	1.07	0.33	2.47	0.71	0.19	1.81	3.07	0.75	8.30
Collateral	5.00	4.23	5.90	3.45	2.84	4.16	8.14	6.49	10.16
$N = 400, T = 5, \psi_0 = 0.2$									
Junior	14.06	10.16	18.33	9.86	6.63	13.62	39.85	28.29	51.51
Jr. Mezz	5.36	3.25	8.02	3.92	2.23	6.18	18.11	10.41	27.64
Sr. Mezz	2.54	1.32	4.29	1.73	0.83	3.10	8.04	3.73	14.48
Senior	1.07	0.46	2.07	0.71	0.28	1.49	3.07	1.11	6.74
Collateral	5.00	4.32	5.77	3.45	2.90	4.09	8.14	6.73	9.74
$N = 100, T = 20, \psi_0 = 1.0$									
Junior	14.06	7.24	23.47	9.86	4.42	17.81	39.85	19.59	61.80
Jr. Mezz	5.36	1.93	10.90	3.92	1.24	8.62	18.11	5.75	36.71
Sr. Mezz	2.54	0.66	6.19	1.73	0.37	4.68	8.04	1.60	21.61
Senior	1.07	0.17	3.30	0.71	0.09	2.47	3.07	0.34	11.41
Collateral	5.00	3.52	7.05	3.45	2.39	4.88	8.14	5.73	11.27
$N = 400, T = 5, \psi_0 = 1.0$									
Junior	14.06	6.87	28.76	9.86	4.73	20.91	39.85	21.78	64.40
Jr. Mezz	5.36	2.50	11.74	3.92	1.79	8.64	18.11	8.29	35.19
Sr. Mezz	2.54	1.11	5.73	1.73	0.72	3.94	8.04	3.10	18.01
Senior	1.07	0.41	2.51	0.71	0.25	1.74	3.07	0.94	8.05
Collateral	5.00	2.99	8.04	3.45	2.04	5.58	8.14	5.76	11.07

Table 6: 90% confidence intervals for risk metrics of a high-grade CDO.

Tranche	PD (%)			EL (%)			$EL_{0.95}$ (%)		
	True	5th	95th	True	5th	95th	True	5th	95th
$N = 100, T = 20, \psi_0 = 0.2$									
Junior	13.24	7.96	19.30	9.56	5.24	15.10	35.17	20.01	51.20
Jr. Mezz	5.41	2.50	9.49	3.89	1.61	7.35	17.06	7.10	30.59
Sr. Mezz	2.53	0.90	5.24	1.87	0.59	4.13	8.62	2.67	18.81
Senior	1.29	0.35	3.12	0.75	0.17	2.04	3.35	0.70	9.54
Collateral	1.00	0.69	1.41	0.68	0.44	0.99	1.92	1.22	2.85
$N = 400, T = 5, \psi_0 = 0.2$									
Junior	13.24	8.21	19.24	9.56	5.51	14.91	35.17	21.50	50.07
Jr. Mezz	5.41	2.74	9.24	3.89	1.82	7.07	17.06	8.18	29.19
Sr. Mezz	2.53	1.06	4.96	1.87	0.72	3.87	8.62	3.26	17.54
Senior	1.29	0.45	2.88	0.75	0.22	1.84	3.35	0.93	8.57
Collateral	1.00	0.72	1.36	0.68	0.45	0.97	1.92	1.29	2.76
$N = 100, T = 20, \psi_0 = 1.0$									
Junior	13.24	6.03	23.73	9.56	3.83	18.60	35.17	14.47	58.45
Jr. Mezz	5.41	1.64	11.75	3.89	0.96	9.30	17.06	4.10	37.41
Sr. Mezz	2.53	0.47	6.82	1.87	0.27	5.55	8.62	1.21	24.90
Senior	1.29	0.14	4.34	0.75	0.05	3.02	3.35	0.22	14.32
Collateral	1.00	0.56	1.71	0.68	0.36	1.19	1.92	0.99	3.43
$N = 400, T = 5, \psi_0 = 1.0$									
Junior	13.24	5.27	29.14	9.56	3.69	22.49	35.17	16.10	60.53
Jr. Mezz	5.41	1.93	13.63	3.89	1.32	10.22	17.06	5.95	36.74
Sr. Mezz	2.53	0.78	6.96	1.87	0.53	5.33	8.62	2.33	22.48
Senior	1.29	0.33	3.89	0.75	0.16	2.42	3.35	0.63	11.08
Collateral	1.00	0.48	1.95	0.68	0.31	1.35	1.92	1.03	3.24

Figure 7: Distribution of estimated mezzanine CDO risk metrics given a long data panel with well observed credit factors ($N = 100, T = 20, \psi = 0.2$).

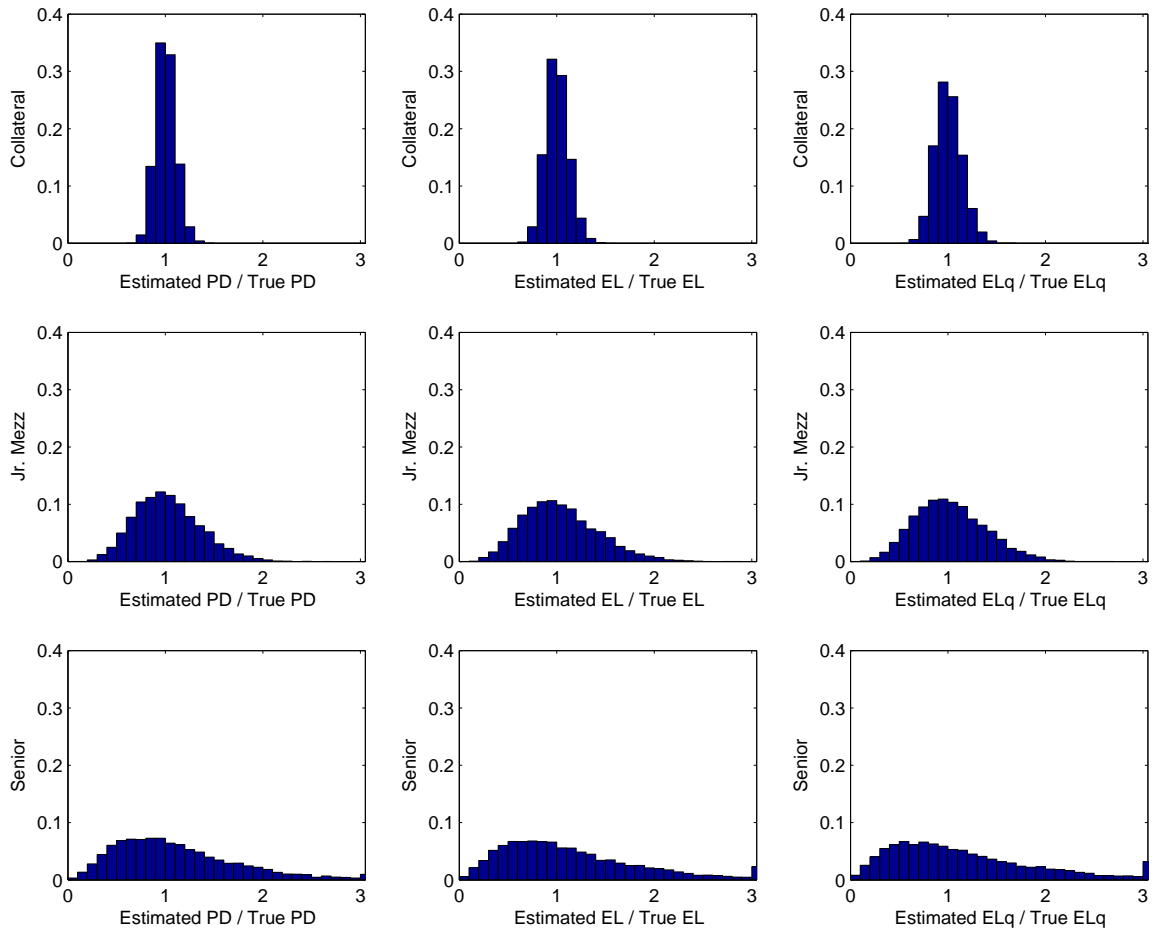


Figure 8: Distribution of estimated mezzanine CDO risk metrics given a wide data panel with well observed credit factors ($N = 400$, $T = 5$, $\psi = 0.2$).

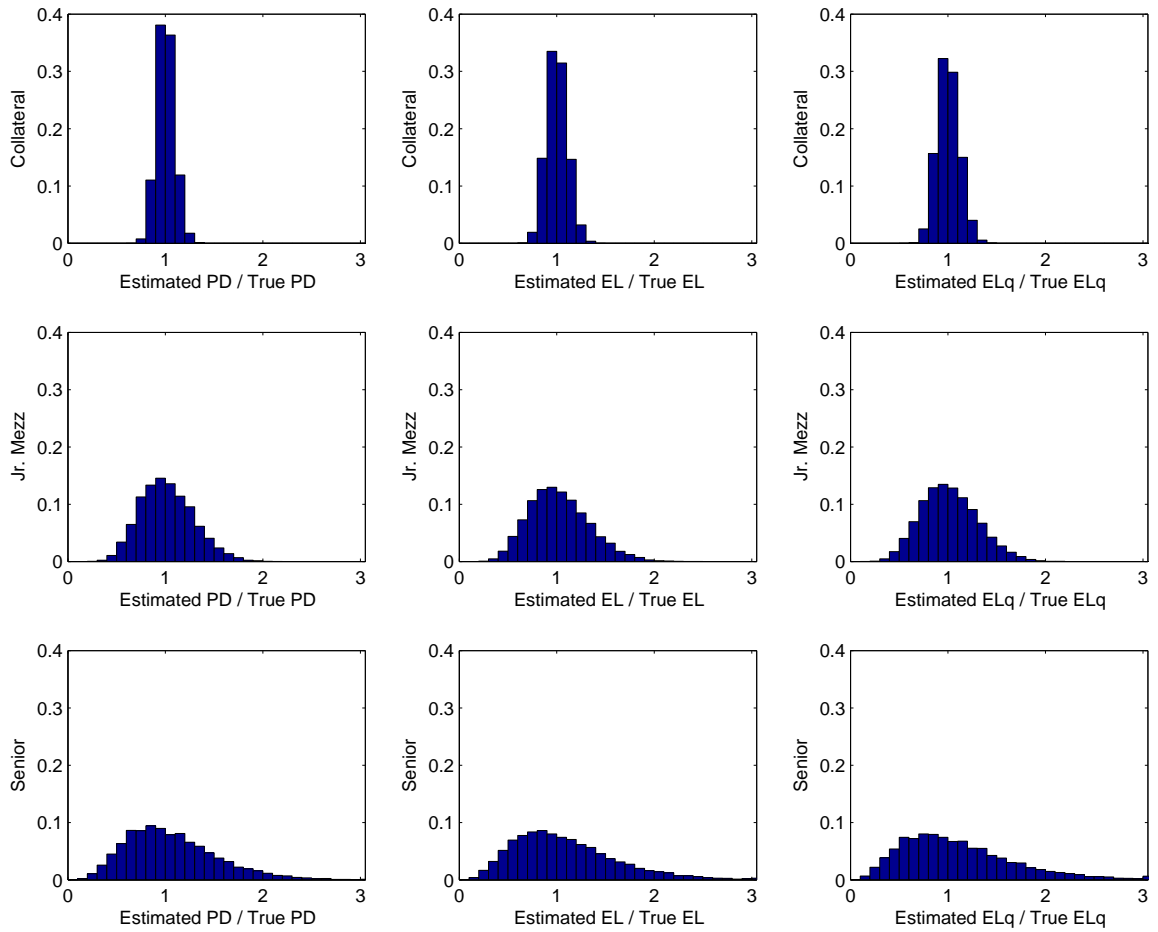


Figure 9: Distribution of estimated mezzanine CDO risk metrics given a long data panel with unobserved credit factors ($N = 100, T = 20, \psi = 1.0$).

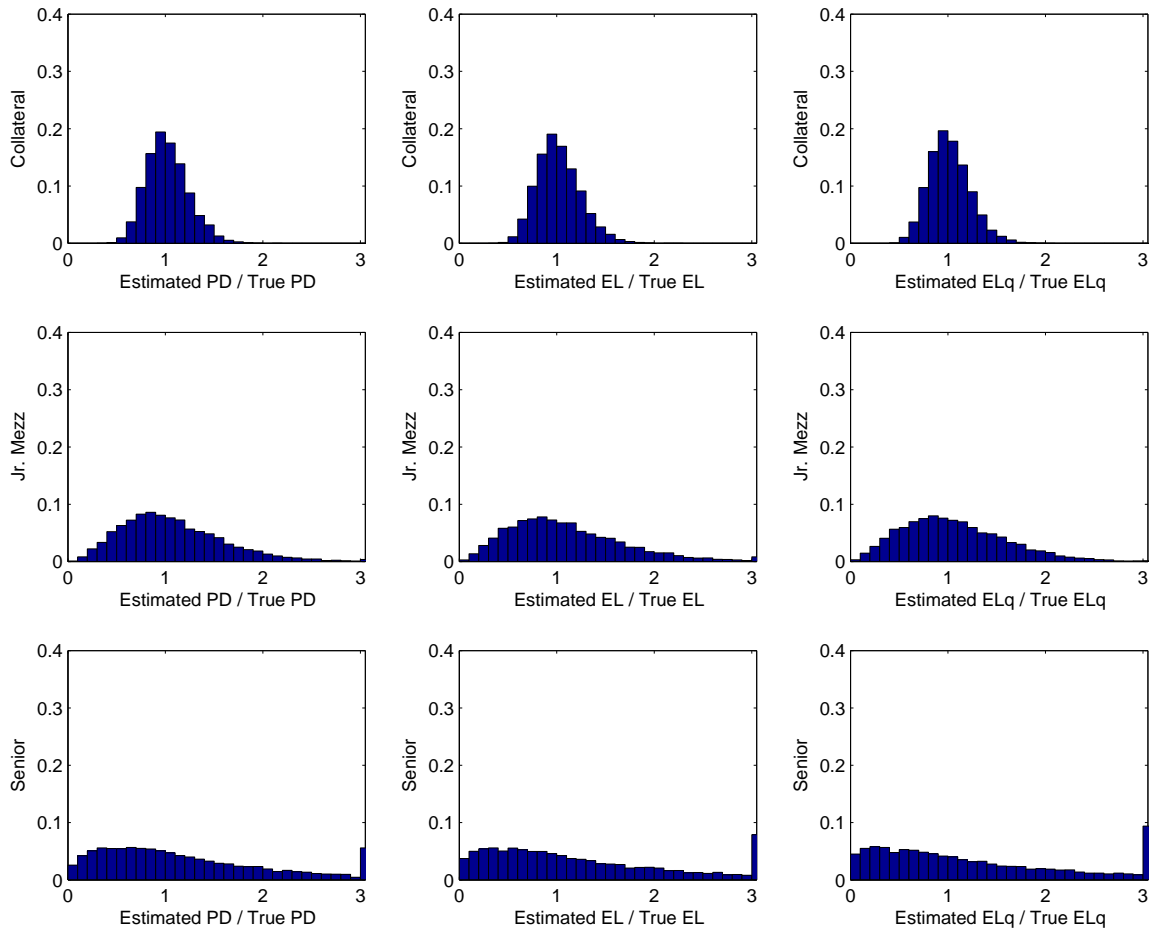


Figure 10: Distribution of estimated mezzanine CDO risk metrics given a wide data panel with unobserved credit factors ($N = 400, T = 5, \psi = 1.0$).

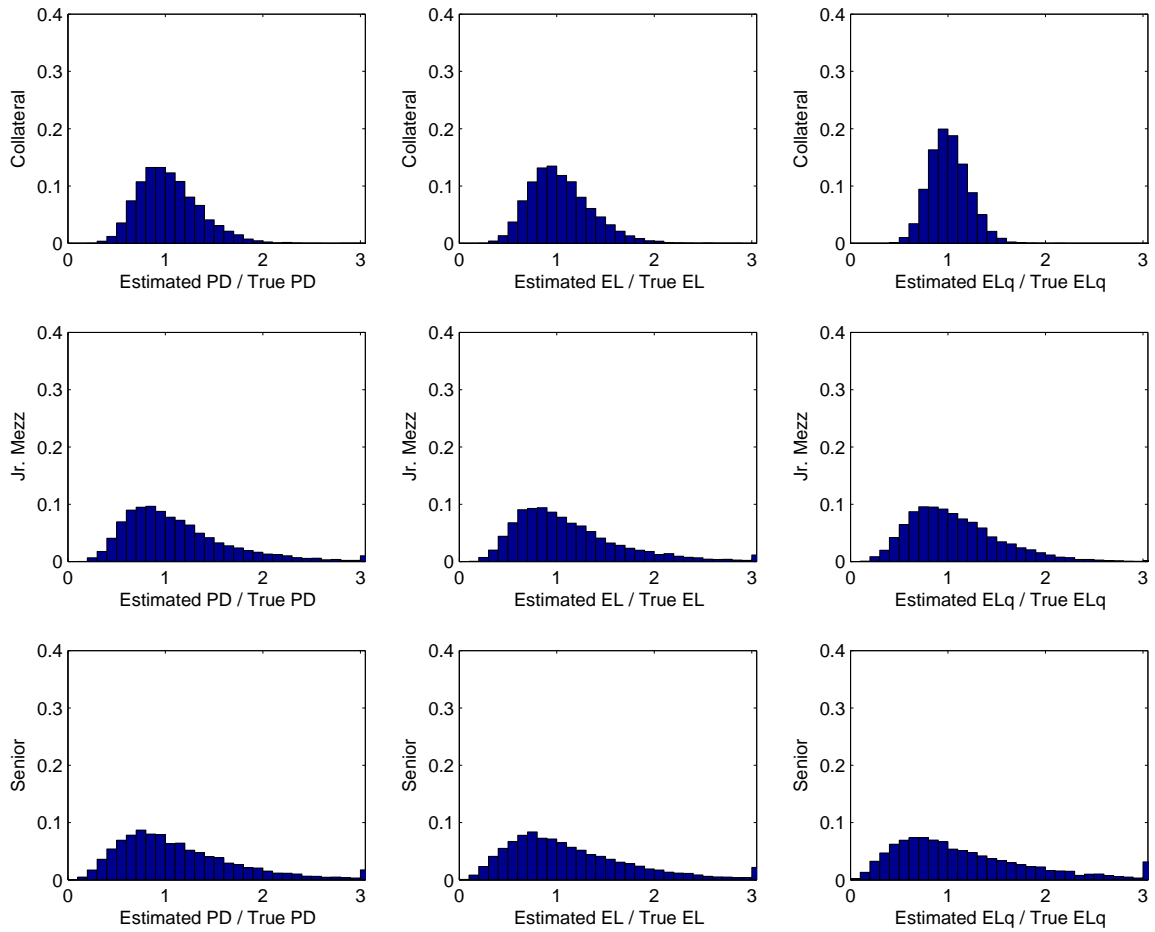


Figure 11: Distribution of estimated high-grade CDO risk metrics given a long data panel with well observed credit factors ($N = 100, T = 20, \psi = 0.2$).

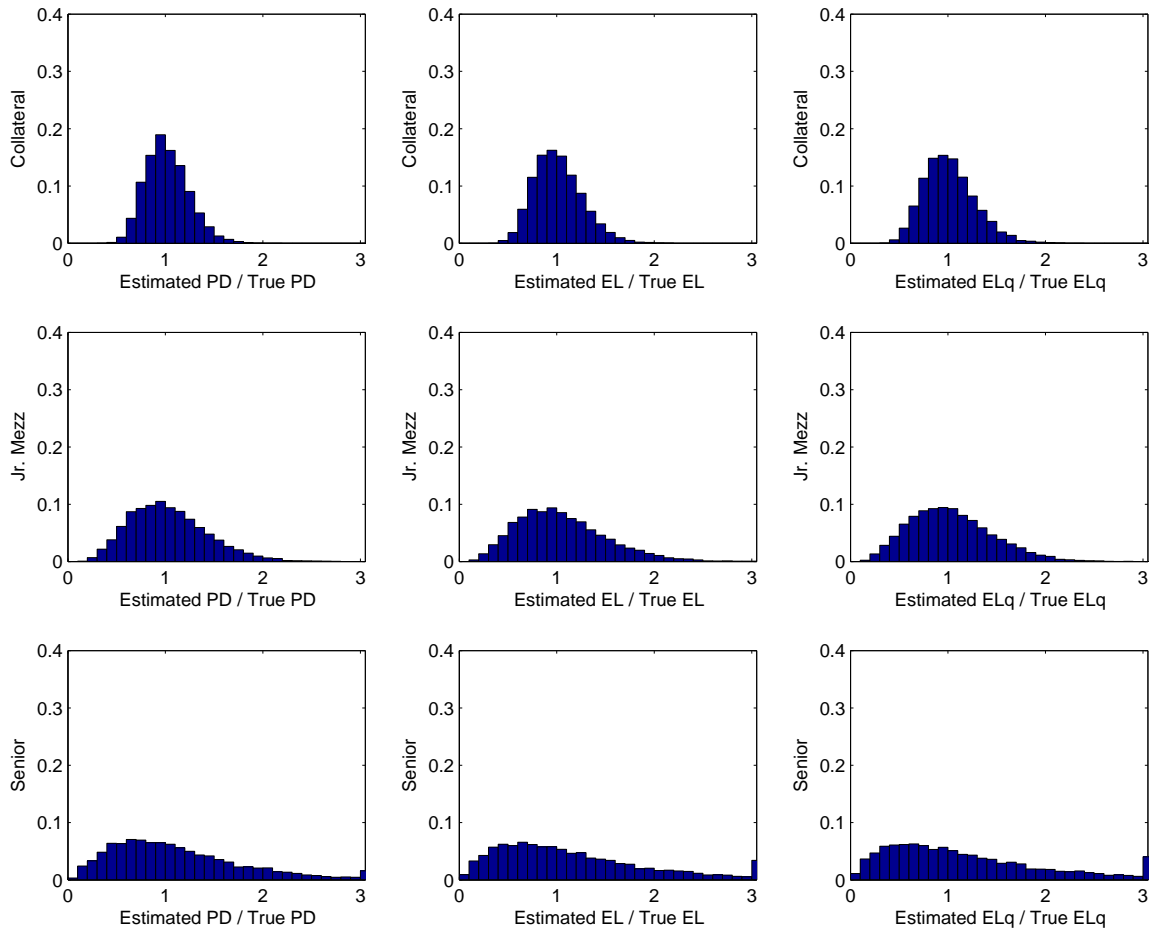


Figure 12: Distribution of estimated high-grade CDO risk metrics given a wide data panel with well observed credit factors ($N = 400, T = 5, \psi = 0.2$).

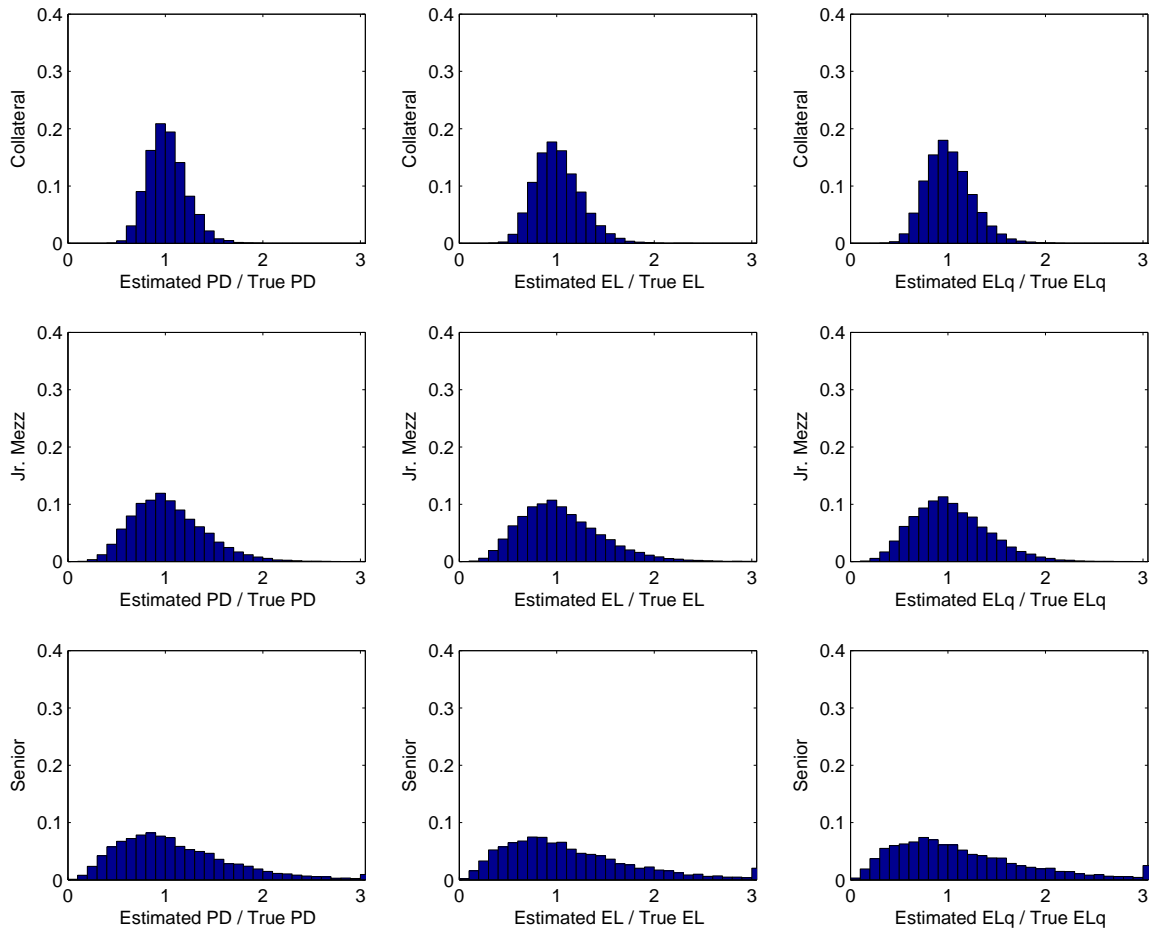


Figure 13: Distribution of estimated high-grade CDO risk metrics given a long data panel with unobserved credit factors ($N = 100, T = 20, \psi = 1.0$).

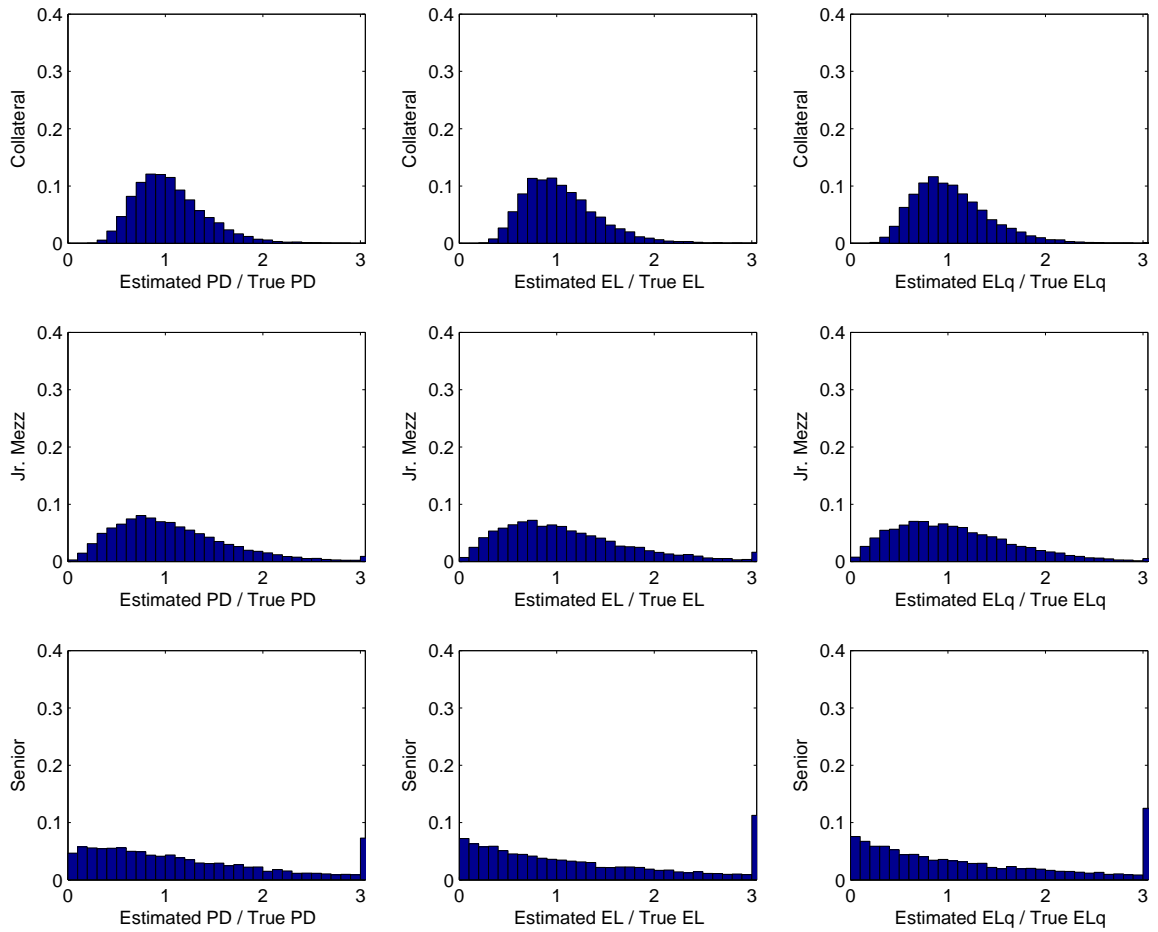
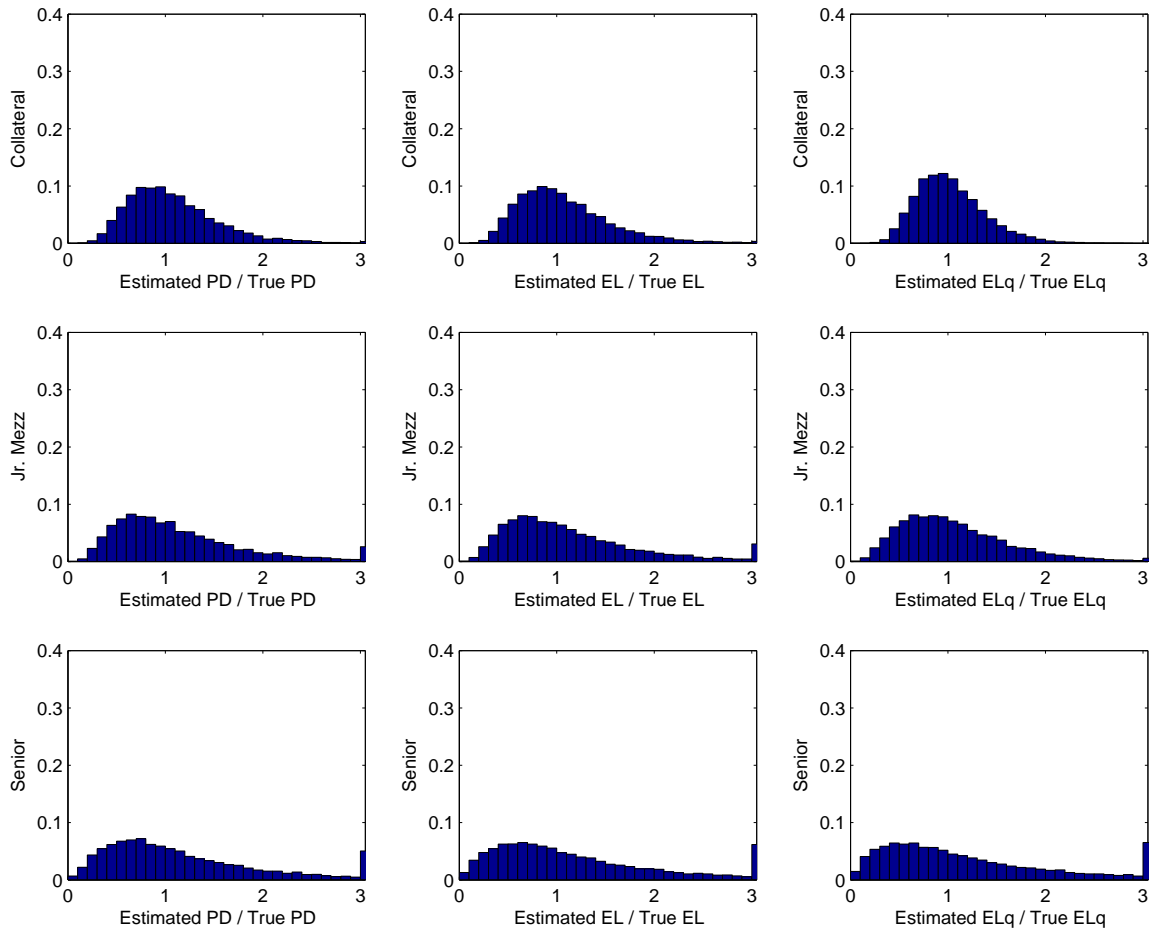


Figure 14: Distribution of estimated high-grade CDO risk metrics given a wide data panel with unobserved credit factors ($N = 400, T = 5, \psi = 1.0$).



The wide dispersion of risk metrics for more senior CDO tranches is striking, particularly when compared with the dispersion of risk metrics for simple bonds with similar true default probabilities and expected losses. For example, the lowest mean-squared-errors for the high-grade CDO deal are produced by the wide data panel with partially observable credit factors. In this setting the simulated 90 percent confidence interval of estimated PDs for the “Senior” tranche spans a range from less than half of the tranche’s true default probability to more than double the true value. In contrast, the simulated confidence interval for a high-grade simple bond with about the same true default probability spans a range of plus or minus only about one-third of the true value. A similar pattern holds for all DGPs considered.

All else equal, the higher a tranche’s position in a CDO deal’s capital structure, the greater is the dispersion in estimated risk metrics relative to the tranche’s true risk metrics. As shown in Section 4, implied risk metrics for more senior tranches are more sensitive to model parameters, so the effects of parameter estimator sampling errors are magnified for these tranches. For all but the most junior CDO tranches, estimated risk metrics are biased upward, reflecting the nonlinear relationship between normal copula/beta model parameters and implied risk metrics. Note however, that the medians of simulated risk metrics (not reported) closely match the true values for all tranches.

The type of collateral backing a CDO deal also affects the accuracy of estimated risk metrics. Comparing results for like CDO tranches backed by high-grade and mezzanine collateral, we see that risk metrics for CDO tranches backed by higher quality bonds have larger sampling errors than those of similar tranches backed by lower quality collateral. As shown in Section 5, when defaults are less common it is more difficult to accurately estimate default probabilities, factor correlations, and parameters describing the distribution of loss rates on defaulted bonds. The greater uncertainty associated with θ results in greater dispersion in estimated risk metrics.

Obviously the quantity and character of available historical data affects the accuracy of estimated risk metrics. As one would expect, observing more cohorts (higher T) or more bonds per cohort (higher N) reduces the dispersion of estimated risk metrics. Holding the total number of data points ($N \cdot T$) fixed, wide, short panels (high N , low T) appear to produce somewhat more accurate risk metrics for structured bonds than long, narrow panels (low N , high T). Proxy data for latent credit factors can dramatically improve the accuracy of risk metrics for both unstructured and structured securities. Notably, such data significantly improve the accuracy of PD and EL estimates for unstructured securities, even though these credit risk metrics do not depend on asset correlation parameters.

7 Conclusions

Overall, the results presented in Section 6 suggest that statements about the credit quality of structured securities should be viewed with considerably more skepticism than comparable statements about the credit quality of unstructured bonds. Because the credit performance of a CDO deal's various debt tranches depend on the full distribution of collateral losses, evaluating the credit quality of CDO notes is inherently more difficult than evaluating the credit quality of unstructured notes with similar underlying risk characteristics. All else equal, CDO notes with greater seniority are more sensitive to model parameters that describe the distribution of collateral losses. Even small errors in estimating these parameters can have significant effects on measures of credit risk for senior CDO notes. Typically it is more difficult to estimate such parameters when collateral defaults are rare. Thus, the uncertainty associated with standard measures of CDO credit risk will be greatest when the note of interest is relatively senior and/or the CDO collateral is of relatively high quality. In other words, the safest CDO notes are likely to be the most difficult to accurately evaluate.

There are strong reasons to believe that in real world settings estimated risk metrics for CDO tranches would be even less accurate than the simulations presented here suggest. The normal copula model used in this analysis is substantially more parsimonious than those used in applied settings. It assumes that CDO collateral is homogeneous with respect to model parameters and imposes a very simple dependence structure on collateral losses. In practice CDO collateral may consist of a diverse mix of securities, factor correlations may vary depending on a security's type, region, or industry sector, and loss severities are likely to be correlated with default rates. Thus, risk metrics for real world CDO tranches typically depend on a great many more parameters than the four described here, and all parameters must be estimated from limited historical data. This analysis also assumes that the collateral loss model used to compute risk metrics is correctly specified and that model parameters are calibrated using minimum variance unbiased estimators. Deviation from either of these assumptions could lead to bias or additional dispersion in estimated risk metrics. Finally, it should be noted that asset-backed-security CDOs (ABS CDOs), such as those that have performed most poorly in recent months, are considerably more complex than the simple bond backed CDO's considered here. ABS CDO credit models must account for the effects of structural credit enhancement on collateral performance and, because ABS CDO tranche payouts depend on complex cash-flow allocation schemes, estimated risk metrics are likely to be particularly sensitive to assumptions about the timing of collateral losses.

Practitioners can deal with parameter uncertainty in one of two ways. Under a "classi-

cal” approach, analysts estimate risk metrics for CDO notes and other risk exposures using the best data and most efficient statistical techniques available, but they also compute confidence intervals and other descriptors of the accuracy of those risk metrics. Risk-management applications that account for parameter uncertainty then involve a two-step process. In the first step, point estimates of relevant risk metrics are taken at face value and used in the usual way. In the second step, analysts examine how their conclusions change under adverse parameter “stress scenarios” determined by the sampling distribution of the parameters used in the first step. Goessl (2005) describes an alternative, Bayesian, approach under which parameter uncertainty is embedded in the process of computing risk metrics. A distribution of model parameters is assumed *ex ante* which is updated as new data become available. Risk metrics of interest are computed by taking expectations that account for both the uncertainty associated with predicting future credit performance and the uncertainty associated with model parameters. Both the classical and Bayesian approaches can be expected to produce qualitatively similar results. Structured credit exposures ought to attract higher economic capital requirements and command higher spreads than would be the case if model parameters were known with certainty, and the relative magnitude of these differences should be greater for safer CDO notes.

These findings lend support to financial regulators’ argument that credit rating agencies should better distinguish between structured and unstructured bond ratings and disclose relevant information about the uncertainties associated with ratings. Moody’s, Standard and Poor’s, and Fitch all currently use the same ordinal rating scales to evaluate traditional corporate debt securities and structured securities, and rating agencies have publicly asserted that bond ratings should be treated as comparable across debt classes. For example, S&P (Wong, Gillis, and Michaux 2007) states:

Our ratings represent a uniform measure of credit quality globally and across all types of debt instruments. In other words, an ‘AAA’ rated corporate bond should exhibit the same degree of credit quality as an ‘AAA’ rated securitized debt issue.

The analysis presented here suggests that the uncertainty associated with ratings for high quality structured products is likely to be much greater than that associated with high-quality corporate debt securities. Quantitative models such as those used by rating agencies to evaluate CDOs may provide useful tools for ranking risks across similar types of debt securities, but, since underlying model parameters are estimated with error, there is a good chance that these models will significantly misstate default probabilities, expected losses, and other measures of the riskiness of these securities.

APPENDIX: Derivation of DGP likelihood

Conditional on X , Y^* and Y have the joint distribution,

$$\begin{bmatrix} Y^* \\ Y \end{bmatrix} | X \sim \mathcal{N} \left(\begin{bmatrix} \sqrt{\omega}\sqrt{1-\psi}\sqrt{\rho}X \\ \sqrt{\rho}X \end{bmatrix}, \begin{bmatrix} \omega S & \sqrt{\omega}(1-\rho)\sqrt{1-\psi} \\ \sqrt{\omega}(1-\rho)\sqrt{1-\psi} & 1-\rho \end{bmatrix} \right).$$

where $S = (1-\psi)(1-\rho) + \psi$. This implies that conditional on Y^* and X , Y is distributed

$$Y | Y^*, X \sim \mathcal{N} \left(\frac{\psi\sqrt{\rho}X + (1-\rho)\sqrt{1-\psi}(Y^*/\sqrt{\omega})}{S}, \frac{\psi(1-\rho)}{S} \right).$$

Thus, we can write the joint distribution of Y^* and Y conditional on X as

$$f(y^*, y|X) = \phi \left(\frac{(y^*/\sqrt{\omega}) - \sqrt{1-\psi}\sqrt{\rho}X}{\sqrt{S}} \right) \phi \left(\frac{Sy - \psi\sqrt{\rho}X - (1-\rho)\sqrt{1-\psi}(y^*\sqrt{\omega})}{\sqrt{S}\sqrt{\psi(1-\rho)}} \right). \quad (7)$$

The first right-hand term is the PDF of Y^* given X . The second term is the PDF of Y given Y^* and X .

Recall that D is an indicator variable that is equal to one if $Y \leq \Phi^{-1}(\pi)$ and zero otherwise. Y^* is only observable when $D = 0$. If $D = 0$ the joint likelihood of Y^* and D given X is

$$f(y^*, D = 0|X) = \int_{\Phi^{-1}(\pi)}^{\infty} f(y^*, y|X) dy = \phi \left(\frac{(y^*/\sqrt{\omega}) - \sqrt{1-\psi}\sqrt{\rho}X}{\sqrt{S}} \right) \times \left(1 - \Phi \left(\frac{S\Phi^{-1}(\pi) - \psi\sqrt{\rho}X - (1-\rho)\sqrt{1-\psi}(y^*/\sqrt{\omega})}{\sqrt{S}\sqrt{\psi(1-\rho)}} \right) \right) \quad (8)$$

The second equality follows directly from (7). If $D = 1$, Y^* is not observable, and the likelihood is

$$f(D = 1|X) = \int_{-\infty}^{\Phi^{-1}(\pi)} f(y|X) dy = \Phi \left(\frac{\Phi^{-1}(\pi) - \sqrt{\rho}X}{\sqrt{1-\rho}} \right). \quad (9)$$

Combining (8) and (9) yields (5).

References

- ANDERSEN, L., AND J. SIDENIUS (2005a): “CDO Pricing with Factor Models: Survey and Comments,” *Journal of Credit Risk*, 1(3), 71–88.
- (2005b): “Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings,” *Journal of Credit Risk*, 1(1), 29–70.
- BASEL COMMITTEE ON BANKING SUPERVISION (2004): *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Bank for International Settlements, Basel.
- BURTSCHHELL, X., J. GREGORY, AND J.-P. LAURENT (2007): “Beyond the Gaussian Copula: Stochastic and Local Correlaiton,” *Journal of Credit Risk*, 3(1).
- DULLMANN, K., J. KULL, AND M. KUNISCH (2008): “Estimating Asset Correlations from Stock Prices or Default Rates – Which Method is Superior?,” *Deutsch Bundesbank Discussion Paper 04/2008*.
- FENDER, I., N. TARASHEV, AND H. ZHU (2008): “Credit Fundamentals, Ratings and Value-at-Risk: CDOs Versus Corporate Exposures,” *BIS Quarterly Review*, pp. 87–101.
- FREY, R., AND A. MCNEIL (2003): “Dependent Defaults in Models of Portfolio Credit Risk,” *Journal of Risk*, 6, 59–92.
- FREY, R., A. MCNEIL, AND M. NYFELER (2001): “Copulas and Credit Models,” *Risk*, pp. 111–114.
- FU, Y., J. GLUCK, P. MAZATAUD, AND D. ROSA (2004): “Moody’s Revisits Its Assumptions Regarding Corporate Default (and Asset) Correlations for CDOs,” Discussion paper, Moody’s Investors Services.
- GILKES, K. (2004): *CDO Evaluator Version 3.0: Technical Document*. Standard and Poor’s, Structured Finance Criteria Report.
- GOESSL, 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, 199–232.
- GORDY, M. B., AND E. A. HEITFIELD (2002): “Estimating Default Correlations from Short Panels of Credit Rating Performacne Data,” Working Paper.
- GUPTON, G., C. FINGER, AND M. BHATIA (1997): *CreditMetrics – Technical Document*. The RiskMetrics Group.
- HAMERLE, A., AND D. ROSCH (2005): “Misspecified Copulas in Credit Risk Models: How Good is Gaussian?,” *Journal of Risk*, 8(1).

- HANSEN, S., AND T. SCHUERMAN (2006): “Confidence Intervals for Probabilities of Default,” *Journal of Banking and Finance*, 30, 2281–2301.
- JOLIVET, G., L. LASSALVY, M. MUELLER-HEUMANN, AND J. SIELER (2006): *CDOROM 2.3 Users Guide*. Moody’s Investors Services.
- KEALHOFER, S. (1998): “Portfolio Management of Default Risk,” Discussion paper, KMV Corporation.
- KOO, M., J. CROMARTIE, AND P. VEDOVA (2006): “Global Rating Criteria for Collateralised Debt Obligations,” Discussion paper, DerivativeFitch, Credit Products Criteria Report.
- LI, D. X. (2000): “On Default Correlations: A Copula Function Approach,” *Journal of Fixed Income*, 9, 43–54.
- LOFFLER, G. (2001): “The Effects of Estimation Error on Measures of Portfolio Credit Risk,” *Journal of Banking and Finance*, 27, 1427–1453.
- MERTON, R. C. (1974): “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *Journal of Finance*, 29, 449–470.
- PRESIDENT’S WORKING GROUP ON FINANCIAL MARKETS (2008): “Policy Statement on Financial Market Developments,” Policy statement.
- PYKHTIN, M., AND A. DEV (2002): “Credit Risk in Asset Securitisations: an Analytical Model,” *Risk*, pp. S16–S20.
- TARASHEV, N., AND H. ZHU (2008): “Specification and Calibration Errors in Measures of Portfolio Credit Risk: The Case of the ASRF Model,” *International Journal of Central Banking*, Forthcoming.
- TOUTAIN, O., D. ROSA, Y. FU, P. MAZATAUD, AND G. JOLIVET (2005): “Moody’s Revisits Its Assumptions Regarding Structured Finance Default (and Asset) Correlations for CDOs,” Discussion paper, Moody’s Investors Services.
- WONG, C. R., T. G. GILLIS, AND F. MICHAUX (2007): “Principles-Based Rating Methodology for Global Structured Finance Securities,” Discussion paper, Standard and Poor’s.