

MODELING METHODOLOGY

Economic Capital Model Validation: A Comparative Study

Authors

Zhenya Hu
Amnon Levy
Jing Zhang

Contact Us

Americas
+1-212-553-1653
clientservices@moodys.com

Europe
+44.20.7772.5454
clientservices.emea@moodys.com

Asia (Excluding Japan)
+85 2 2916 1121
clientservices.asia@moodys.com

Japan
+81 3 5408 4100
clientservices.japan@moodys.com

(A version of this paper appears in the *Journal of Risk Model Validation*, March 2013)

Abstract

Evaluating an economic capital model's quality is an important issue for both financial institutions and regulators. However, a major obstacle to economic capital model validation remains: the limited number of events where extreme losses were realized. Credit risk models normally estimate economic capital over a one-year horizon and at a high confidence interval, which requires a large number of years to produce sufficient observations to evaluate model accuracy.

Using a long history of public firm defaults from Moody's Investor Services and Moody's Analytics, this study illustrates a validation approach for jointly testing the impact of PD and correlation upon model performance. We construct predicted default distributions using a variety of PD and correlation inputs and examine how the predicted distribution compares with the realized distribution. The comparison is done by looking at the percentile of realized defaults with respect to the predicted default distribution. We compare the performance of two typical portfolio parameterizations: (1) a through-the-cycle style parameterization using agency ratings-based long-term average default rates and Basel II correlations; and (2) a point-in-time style parameterization using public EDF credit measure, and Moody's Analytics Global Correlation Model (GCorr™).

Results demonstrate that a through-the-cycle style parameterization results in a less conservative view of economic capital and substantial serial correlation in capital estimates. Results also show that when point-in-time measures are used, the tested economic capital model produces consistent and conservative economic capital estimates over time.

Table of Contents

1	Introduction.....	5
2	Methodology.....	6
3	Data.....	9
	3.1Portfolio Description.....	12
4	Portfolio Analysis.....	15
	4.1Through-the-Cycle Case.....	15
	4.2Point-in-Time Case.....	17
5	Conclusion.....	19
	References.....	21

1 Introduction

Evaluating economic capital models for credit risk is important for both financial institutions and regulators. However, a major impediment to model validation remains limited data in the time series due to the following issues. First, defaults are rare events. One can only observe a limited number of events where extreme losses were realized; second, credit risk models typically estimate value distribution over a one-year horizon, and the number of non-overlapping years is limited in availability. This problem differs from the analysis of market risk, which focuses on a much shorter horizon.

The data constraint makes it difficult to transfer the methodology applied to back-testing market risk models to credit risk models, which require a very long time period to produce sufficient observations for reasonable tests of forecast accuracy. To overcome the data limitations, Lopez and Saidenberg (2000) suggested cross-sectional resampling techniques, which resample from the original panel dataset of credits to generate additional credit default/loss data for model evaluation. Frerichs and Löffler (2002) suggested a Berkowitz (2001) procedure to detect mis-specified parameters in asset value models, focusing on asset correlations. The authors conducted Monte Carlo simulations to show that a loss history of ten years can be sufficient to detect mis-specified asset correlation in a two-state credit risk model, but, when applying the test to a multi-state credit risk model, the authors found “incorporating migration and recovery rate uncertainty reduces the test’s power.”¹ BCBS (2009) interpreted validation more expansively as an evaluation of all the processes that provide evidence-based assessment regarding an EC model’s fitness for purpose. The types of validation processes are qualitative processes (e.g., use test, management oversight and data quality checks, etc.) and quantitative processes such as validation of inputs and parameters, benchmarking, back testing, etc. Following these supervisory standards, Jacobs (2010) surveyed the different existing EC model validation practices in the banking industry and illustrated several quantitative approaches (benchmarking, sensitivity analysis, and testing for predictive accuracy) by presenting results of a bank risk aggregation study from Inanoglu, et al (2010).

In general, there has been limited research that either empirically evaluates credit portfolio risk models or that theoretically develops statistical evaluation methods. In its 2009 report, “Range of Practices and Issues in Economic Capital Frameworks,” the Basel Committee on Banking Supervision recognizes that so far proposed statistical tests of portfolio credit models have weak power. “. . . The validation of economic capital models is at a very preliminary stage. . . . The validation techniques are powerful in some areas such as risk sensitivity but not in other areas such as overall absolute accuracy or accuracy in the tail of the loss distribution.”

While these approaches provide insights into credit portfolio model validation, little is known empirically about the performance of a credit portfolio model. Given the challenges, we proceed by focusing on the parts of the distribution for which we have data. We illustrate a validation approach using historical corporate default experience dating back to 1983. This extensive data set is comprised of the defaults associated with the firms rated by Moody’s Investors Service (MIS) and the defaults associated with public firms collected by Moody’s Analytics (MA).² We construct predicted default distributions using different types of PD and correlation inputs, and we then examine how the predicted distribution compares with the realized distribution. We compare by looking at the percentiles of realized defaults with respect to the predicted default distributions. We next look at the performance of two typical portfolio parameterizations: (1) a through-the-cycle style parameterization using agency ratings-based long-term average default rates, and Basel II correlations, and (2) a point-in-time style parameterization using public EDF credit measure, and Moody’s Analytics Global Correlation Model (GCorr).

The remainder of this paper is organized as follows:

- Section 2 presents the validation procedure.
- Section 3 describes the data for portfolio construction and alternative PD and correlation inputs.

¹The two-state credit risk model neglects migration risk and assume zero recovery for all loans. As a consequence, the loss distribution is fully described by the distribution of the number of defaults within a portfolio. The multi-state model incorporates both migration risk and recovery rate uncertainty.

² For Moody’s definition of default, see “Corporate Default and Recovery Rates, 1920–2010” (2011).

- Section 4 discusses a comparison of realized default rates and predicted default rates under two test scenarios.
- Section 5 concludes.

2 Methodology

Our procedure begins with a portfolio of credit exposures of N companies active during the beginning of a year. A credit portfolio model parameterized with default probabilities and correlations implies a distribution of default rates. The corresponding percentile of realized defaults is determined with respect to the estimated distribution of default counts. We estimate a percentile for each year. If the predicted loss distribution is equal to the true one, the obtained percentiles should follow independent and identically distributed uniform distributions.³

Rosenblatt (1952) defines the transformation

$$X_t = \int_{-\infty}^{y_t} \hat{f}(u) du = \hat{F}(y_t) \quad (1)$$

Where y_t is the *ex post* portfolio profit/loss realization and $\hat{f}(\cdot)$ is the *ex ante* forecast loss density. Rosenblatt shows that when applying the estimate cumulative distribution function $\hat{F}(\cdot)$ to observed losses, if the estimated loss distribution is equal to the true one, the transformed variable X_t follows i.i.d. $U(0,1)$.

The detailed procedures follow. We create two sets of equally-weighted, vanilla term loan portfolios consisting of (1) all the MIS rated corporate firms and (2) all the non-financial public firms active in January of the test years. The number of active firms differs across time, so that the composition of the yearly portfolios varies over time for both portfolios. For each name, depending upon the test scenario, we assign a PD value upon either a rating-mapped PD or an EDF credit measure, and we then assign a GCORR correlation or asset correlation used by the Basel II IRB.⁴ All loans have a 0% annual fixed rate coupon LGD of 100% and a one year maturity.

We analyze portfolios in Moody's Analytics RiskFrontier™ to obtain portfolio value distributions at a one year horizon using one million simulation trials. Because the LGD is set to 100% and no coupons are accumulated before horizon, the portfolio value distribution is then converted to a distribution of realized default. For example, suppose a portfolio has 100 instruments at the beginning of the year, each with a \$1 million of commitment amount, i.e., \$100 million in total. With zero defaults, the portfolio value should remain \$100 million. If, on a particular trial, the portfolio value is \$95 million, the associated loss is \$5 million and the number of defaulted obligors is five in this trial. Similarly, one can calculate the number of defaults in other trials and tabulate results from each trial to construct a default distribution.

We next determine which percentile of the predicted distribution is associated with the realized default. Suppose during a given year, the actual number of defaults is 10, which lies at the 95th percentile of the estimated default distribution mentioned above. We record a 95th percentile value for that year. Repeating this exercise for the remaining years, we obtain a time series of mapped percentiles. If the model is correct, every percentile value should occur in $[0,100]$ with the same probability. In other words, the percentiles should be independent and identically uniformly distributed. In addition, the frequency of exceptions (the number of actual defaults that exceed the number of predicted defaults indicated by VaR estimate) should be in-line with the selected target level of confidence. For example, for VaR calculated at 99% confidence interval, exceptions can be expected every 100 mapped percentiles.

The following figure provides a visual representation of the mapping process. Each dot “•” in the right hand plot represents the mapped percentile of realized defaults with respect to the model-implied distribution at the specified year. The model-implied distribution varies each year due to the change in portfolio composition and the changes in PD and correlation.

³ See Rosenblatt (1952) and Berkowitz (2001) for details of the Rosenblatt transformation.

⁴ The brief description of these two correlation measures can be found in Section 3.

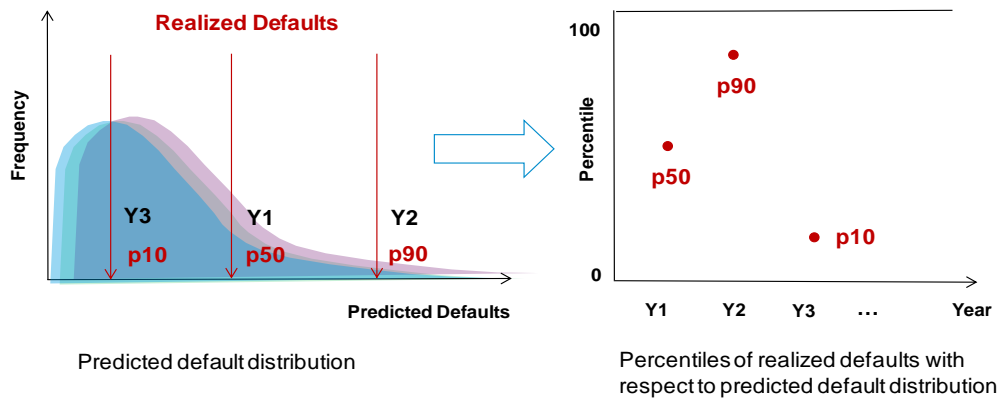


Figure 1 Realized portfolio loss percentiles.

Figure 2 through Figure 5 show examples of inaccurate models of various kinds. In Figure 2, the percentiles are near the top end of the range, suggesting that the model underestimates the probabilities of defaults so that the realized defaults largely fall into the tail region of the estimated distribution. On the contrary, if the model overestimates the probabilities of defaults, the realized defaults tend to concentrate on the low end of the estimated default distribution. The level of mapped percentiles becomes low, as shown in Figure 3.

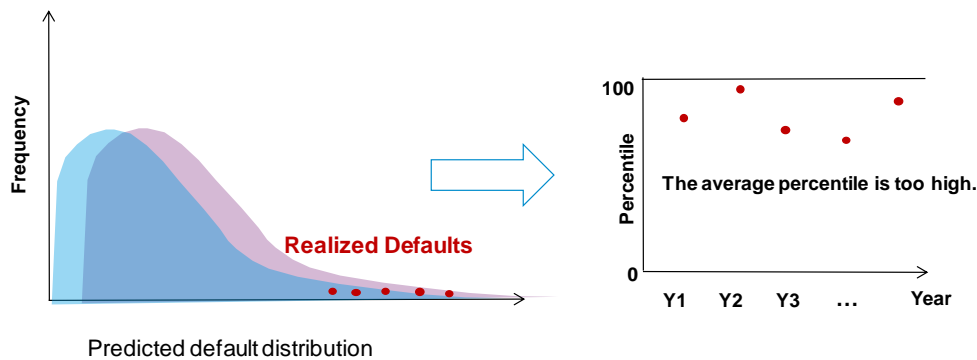


Figure 2 An example of PD underestimation.

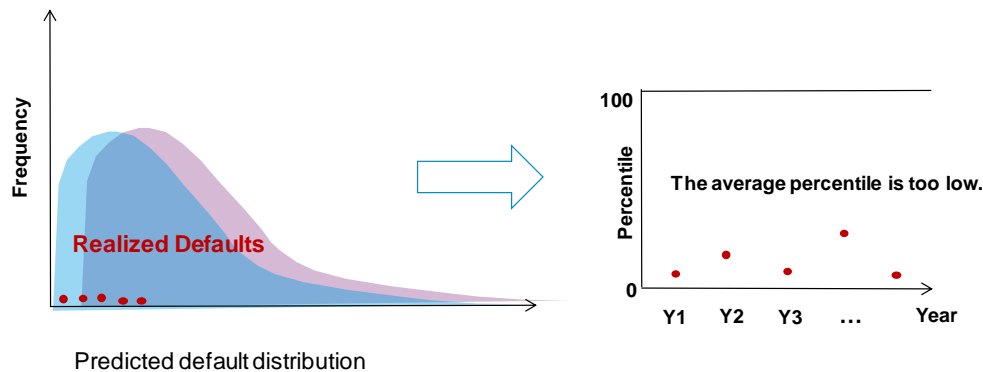


Figure 3 An example of PD overestimation.

We now explore the impact of correlations upon the distribution of percentiles under the assumption that the PD model is correct, and that the average PD is consistent with the average realized default rate. In Figure 4, we see that the percentile values are widely scattered, located around either the top or the bottom of the range. The large dispersion of dots indicates that the model underestimates the correlation between obligors, as the model fails to capture the extreme co-movements corresponding to joint defaults of underlying credits. Figure 5 shows the opposite scenario. The model's correlation overestimation results in realized values falling within a narrow range around 50 percent. The dispersion of percentile values is limited.

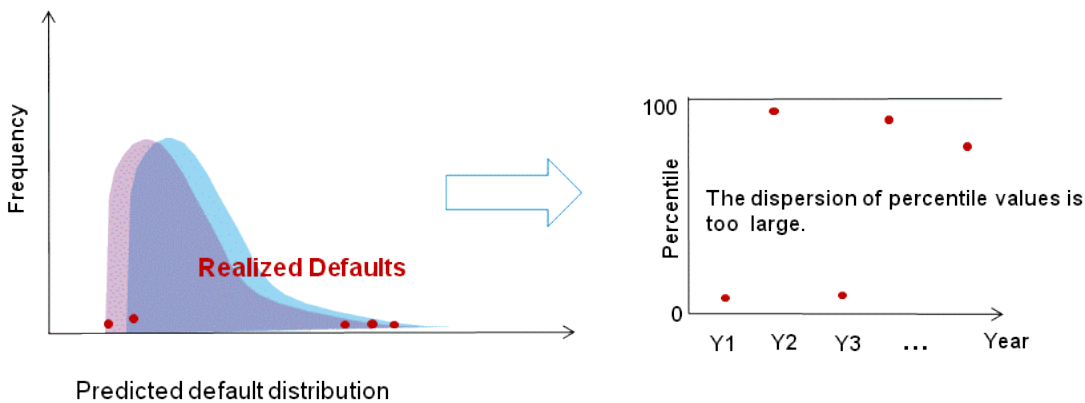


Figure 4 An example of correlation underestimation.

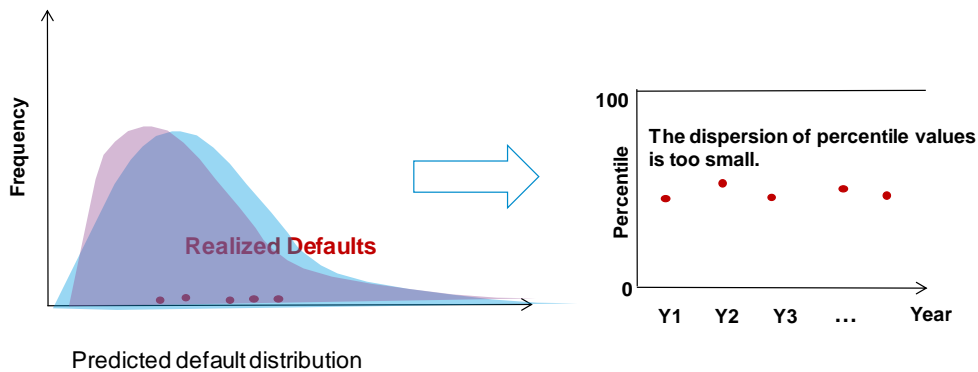


Figure 5 An example of correlation overestimation.

The following plot shows an example of a good model. On average, the realized defaults lie around the 50th percentile of the predicted default distribution. With a 10% of VaR level, 90 percent of dots are expected to lie between the 5th and 95th percentiles.

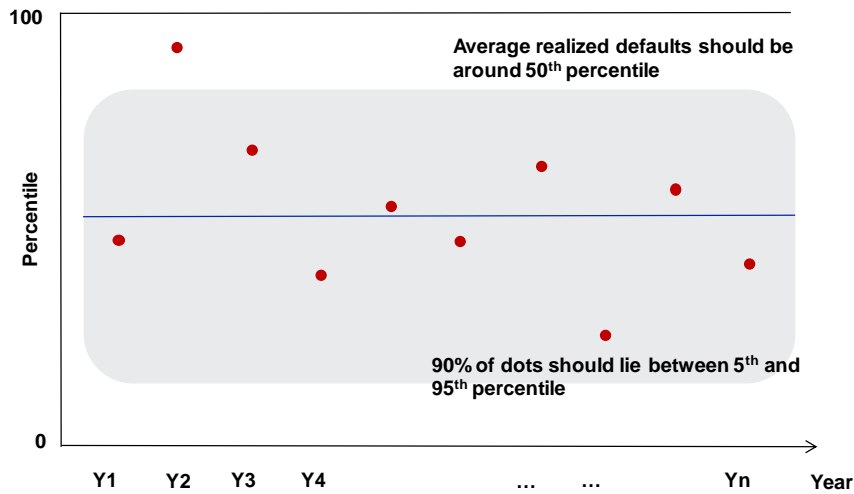


Figure 6 An example of good model.

There are a variety of statistical tests available for accessing the adequacy of VaR measures. Kupiec’s (1995) proportion of failures Test (POF test) is a well known VaR back testing procedure. It examines whether the frequency of exceptions is in line with the selected target level (unconditional coverage property).⁵ The exception is defined as the case where the number of actual defaults exceeds the number of predicted defaults indicated by VaR estimate. If the number of exceptions is less than what the selected confidence level would indicate, the model overestimates risk. On the contrary, too many exceptions signal underestimation of risk. A shortcoming of the POF test is that it does not examine the extent to which the independence property is satisfied. An accurate VaR measure should exhibit both the independence and unconditional coverage property. Christoffersen’s (1998) proposed Markov tests examine whether or not the likelihood of a VaR violation depends upon whether or not the violation occurred during the previous period. Christoffersen and Pelletier further suggest a duration test (2004) using the insight that the time between VaR violations should not exhibit any kind of time dependence. Crnkovic and Drachman (1997) suggest the test on unconditional coverage and independence property to multiple VaR levels instead of a single VaR level, to take the magnitude of the violations into account. Finally, Lopez (1999b) suggested loss function-based approaches.

Unfortunately, the above tests have weak statistical power when using small datasets. Take the POF test as an example. Violations occur rarely (the target probability is usually set small), and therefore testing whether violations form a Bernoulli requires a large sample size. Although a variety of statistic tests exist, these tests are data-intensive and not practical for validating credit portfolio models that describe extreme losses.

3 Data

In this study, we use two default databases: defaults included in the Moody’s Investors Service (MIS) and in the Moody’s Analytics (MA) default database, collected and updated from numerous print and online sources worldwide. For the MIS default database, we focus on defaults associated with Moody’s-rated corporate issuers, given that rated defaults are better documented than unrated defaults. Moody’s uses the senior ratings algorithm

⁵ See Kupiec (1995) for the description of the test.

(SRA) to derive issuer-level ratings and are referred to as estimated senior ratings.⁶ We employ data after 1983 in order to make sure the rated firms have consistent ratings during the sample period. These ratings are used as one of the PD inputs in our tests. The second data source, MA's default database of non-financial public companies 1980 and 2010, is the most extensive public company default database available.

The PD inputs for public companies come from Moody's Analytics EDFTM (Expected Default Frequency) credit measures – probabilities of default for firms with publicly traded equity and published financial statements derived using the Vasicek-Kealhofer (VK) model. This model provides a rich framework that treats equity as a perpetual down-and-out option on the underlying assets of the firm. This framework incorporates five different classes of liabilities: short-term liabilities, long-term liabilities, convertible debt, preferred shares, and common shares. The VK model uses an empirical mapping based on actual default data to get the default probabilities, known as EDF credit measures. Volatility is estimated through a Bayesian approach that combines a comparables analysis with an iterative approach. For an overview of the EDF credit measure, see Crosbie and Bohn (2003).

The following figures show the number of unique firms and the associated defaults by year for the MIS and MA databases, respectively. Both datasets provide global coverage.

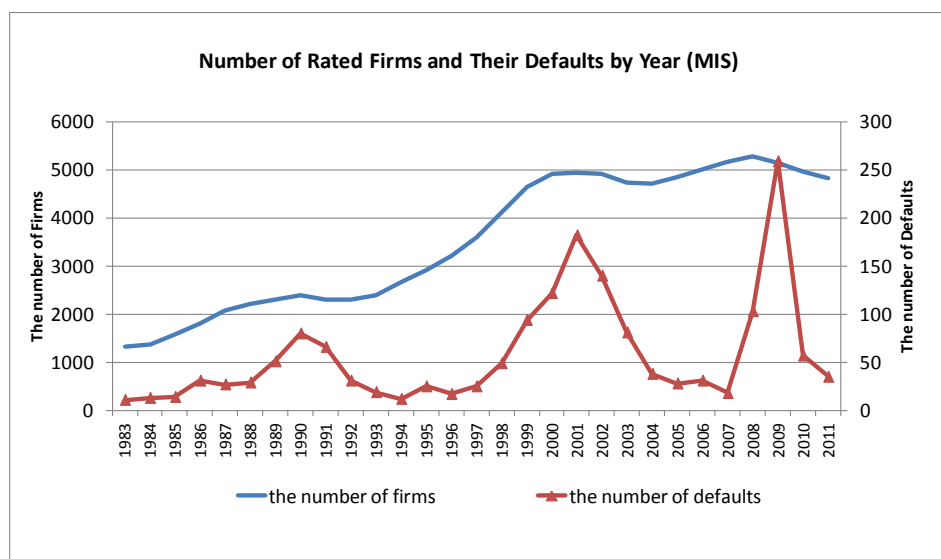


Figure 7 Yearly number of unique rated firms and defaults (from MIS).

⁶ Before 1983, the broad rating category is: Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C. Since April 1982, numerical modifiers (using the 1, 2, 3 modifiers) were appended to each generic rating classification from Aa through Caa. See "Moody's Senior Ratings Algorithm & Estimated Senior Ratings" (2009) for details.

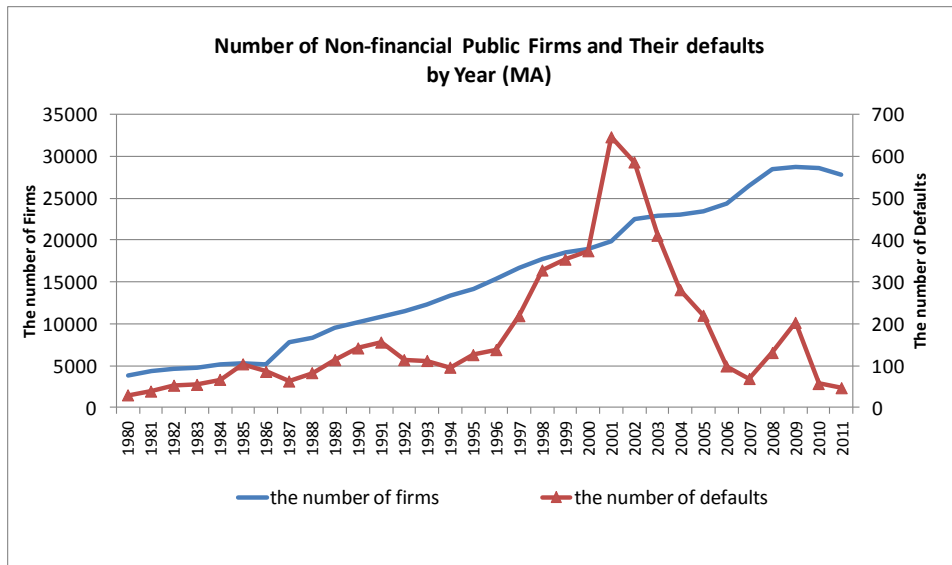


Figure 8 Yearly number of unique non-financial public firms and defaults (from MA).

Note the default databases from MIS and MA used here have different data coverage. If a company has debt rated by MIS but does not have public equity, then its default is not recorded in the MA database.

We use two types of correlation measures.⁷ The first one is defined by the Basel II IRB formula as follows.

$$\rho = a \times \left(\frac{1 - \exp(-c \times PD)}{1 - \exp(-c)} \right) + b \times \left(1 - \frac{1 - \exp(-c \times PD)}{1 - \exp(-c)} \right) \quad (2)$$

Parameters a, b, and c depend upon borrower type. For corporate borrowers, a=0.12, b=0.24, and c=50. The above expression indicates that the asset correlation parameter R is a decreasing function of PD.⁸

The alternative correlation measure is Moody's Analytics global asset correlation (GCorr), a multi-factor model estimated from weekly asset return series. The asset returns are derived from equity returns and liability structure information using an option-theoretic framework. The GCorr model has broad data coverage. For example, it covers more than 34,000 firms in 49 countries and 61 industries, January 2008 through June 2009. First released in 1996, the GCorr model is updated on a regular basis to reflect the most recent dynamics of firms' businesses and industries.

One common default data collection challenge remains the fact that small public companies often disappear without any news or record before they default, or they do not publicly disclose missed payments, both of which creates a number of hidden defaults. To alleviate this hidden defaults problem, in the second case, we restrict the sample to firms above at least \$300 million in annual sales, where we believe hidden defaults are less of an issue. We discuss the hidden default issue further in Section 4.2.

⁷ For details of the Basel II correlation, see the Basel Committee on Banking Supervision, "International Convergence on Capital Measurement and Capital Standards," (2006). For the GCorr model, see Moody's Analytics document "An Overview of Modeling Credit Portfolios," (2008).

⁸ See the Basel Committee on Banking Supervision, (2006), "International Convergence on Capital Measurement and Capital Standards."

3.1 Portfolio Description

We create two portfolios for test purposes. Portfolio-specific settings are highlighted in the following input table in order to make a distinction.

Table 1 Description of Inputs and Settings

	Case 1	Case 2
PD	Long-term average default rate by rating category	EDF
Correlation	Basel II Correlation	GCorr Correlation
Default Data	MIS	MA
Distribution	RiskFrontier	RiskFrontier

We choose these cases as they represent common parameterization approaches by risk management groups at financial institutions. Case 1 represents a regulatory-style parameterization where Basel correlations are used, and default probabilities are based upon a measure frequently associated with a through-the-cycle concept. Meanwhile, Case 2 represents a more point-in-time measure, where both PDs and correlations are parameterized using forward looking measures.

One may notice that if the above two approaches are applied to common datasets the results would be more directly comparable. However, the limited data of public companies with rating histories in our database precludes such an attempt.

3.1.1 Through-the-Cycle Case

In Case 1, we use the long term average of the actual default rate by rating as the PD estimate. As a result, firms with same rating are assigned the same PD value regardless of the sample year. The long term average PD is free of the credit cycle effect and reflects the long-run credit risk level. Table 2 shows the mapping from the rating to long-term average default rate.⁹

Table 2 The Alphanumeric Rating-to-Long-Term Average Default Rate Mapping Table

Rating	Long-Term Average PD	Rating	Long-Term Average PD
Aaa	0.00%	Ba2	0.43%
Aa1	0.00%	Ba3	1.77%
Aa2	0.00%	B1	2.03%
Aa3	0.09%	B2	3.46%
A1	0.06%	B3	5.70%
A2	0.03%	Caa1	7.20%
A3	0.06%	Caa2	10.80%
Baa1	0.15%	Caa3	19.24%

⁹ The long-term average default probability is calculated from raw empirical data and may not increase monotonically as the rating drops.

Baa2	0.17%	Ca	12.96%
Baa3	0.31%	C	9.07%
Ba1	0.56%		

The following figure shows the portfolio composition by letter rating. The composition is relatively stable over time. “Aaa” through “A” rated names account for the largest portion of the whole portfolio most of the time. The “Caa” through “C” rated names represent the smallest portion of the entire portfolio.

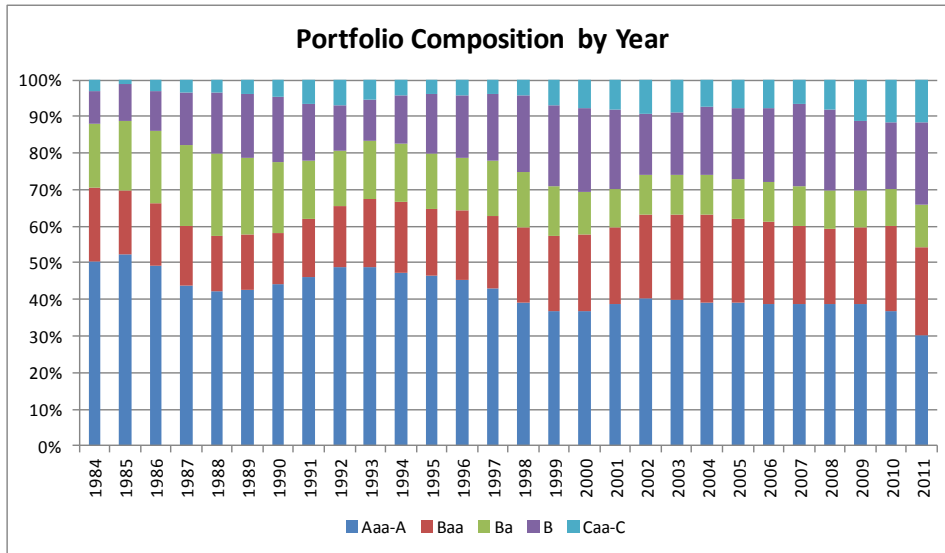


Figure 9 Portfolio composition by rating.

As mentioned previously, we use the long-term average of the realized default rate by rating category as the PD estimate for each name in the portfolio in a given year. Equation (1) is then parameterized with this PD to calculate the corresponding correlation defined in Basel II IRB. Figure 10 shows the average PD and Basel RSQ of portfolios by year.

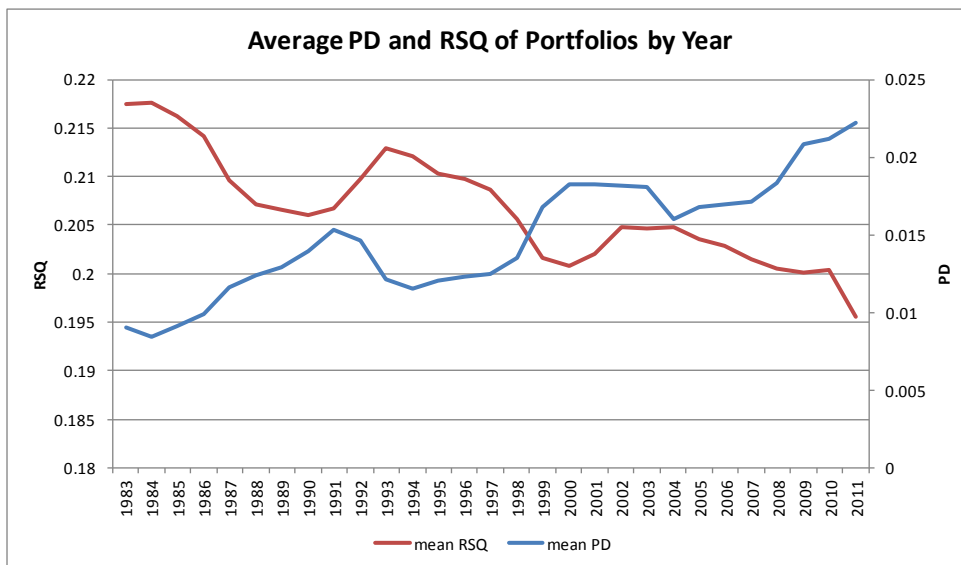


Figure 10 Average PD and RSQ of portfolios by Year (Case 1).

3.1.2 Point-in-Time Case

In Case 2, we construct yearly US portfolios by pooling all US non-financial firms available at the beginning of the year beginning in 1980. We use the MA EDF credit measure as the PD measure and GCorr as the correlation input. GCorr is a multifactor model that assumes co-movements among asset returns are driven by a set of common factors. Unlike historical correlations containing random noise, in addition to useful information, a well-constructed multifactor model produces more accurate forward-looking correlation measures. The GCorr model is updated regularly. Specifically, GCorr1996, GCorr1999, GCorr2002, GCorr2003, GCorr2004, GCorr2005, GCorr2006, GCorr2007, GCorr2009, and GCorr2010 have been released since 1996. In constructing the portfolio, we use the proper version of the GCorr model as of the test year. For portfolios earlier than 1996, we use the modeled R-squared calculator for GCorr2002 to calculate modeled RSQ, as it is the earliest available version. The Modeled R-squared Calculator is a non-linear econometric model of three underlying factors: country weight, industry weights, and firm size.¹⁰

Figures 11 and 12 show the number of firms in the US portfolios.

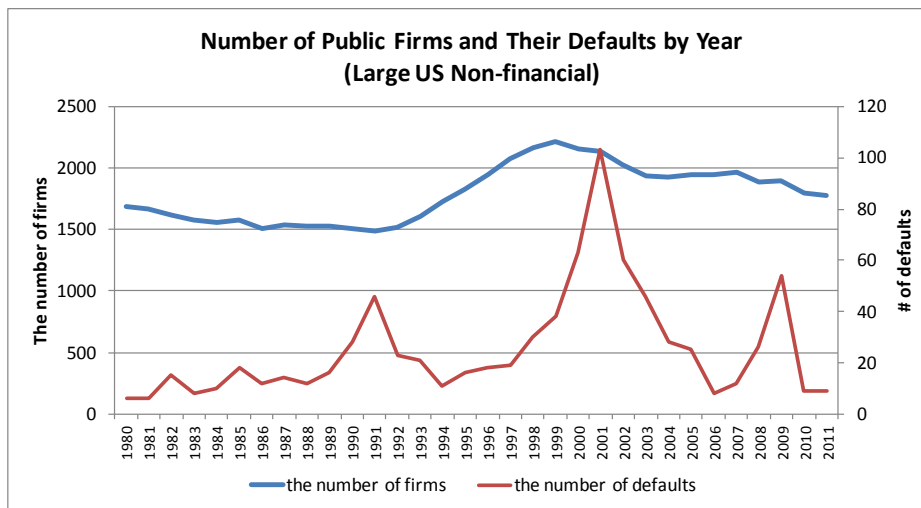


Figure 11 Number of unique firms in portfolios by year (US large non-financials).

Figure 12 shows the average EDF and the GCorr RSQ.

¹⁰ See Moody's Analytics white paper "Moody's KMV Private Firm R-squared Model" for details.

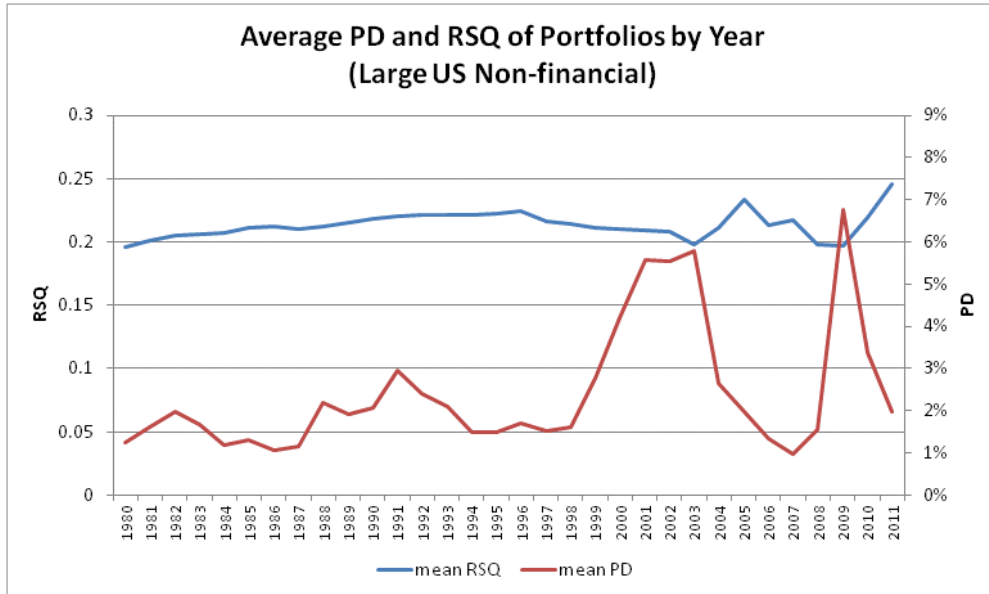


Figure 12 Average GCORR RSQ and EDF1 by year (US large non-financials).

4 Portfolio Analysis

We analyze the yearly portfolios above in RiskFrontier in order to obtain the estimated portfolio value/default distribution at the one year horizon. We then compute the percentiles of realized defaults. Note that the estimated distribution may not be smooth. For example, in 1983, there are 1,330 firms and 11 realized defaults in the sample portfolio, while in the estimated portfolio distribution, the number of 11 defaults corresponds to a range of percentiles from the 60th percentile to the 64th percentile. Under this circumstance, we take an average of 60 and 64 and use the mean value of 62 as the percentile for the default realization.

4.1 Through-the-Cycle Case

We now look at through-the-cycle (Case 1). In Figure 13, each diamond “◆” represents the assigned percentile after mapping the realized number of default to the estimated default distribution at the corresponding year under the Case 1 parameterization. The plot demonstrates a large swing in the percentile points and a strong serial correlation between percentiles.

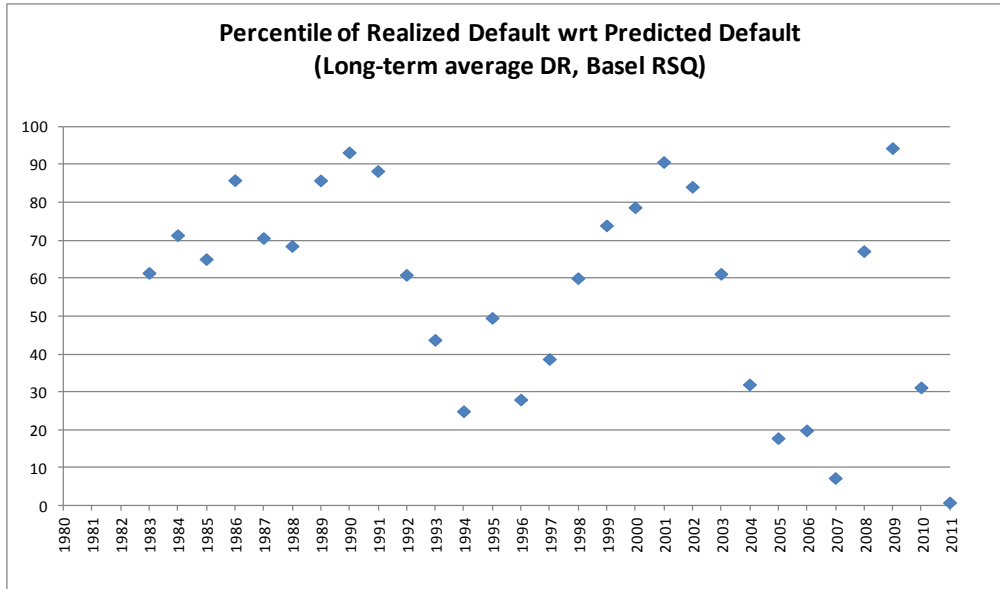


Figure 13 Percentile of realized defaults with respect to RiskFrontier predicted default distribution (Case 1).

In Figure 14, by looking at the autocorrelation function (ACF) plot of the differenced percentile values with a 95% confidence band, we can identify that the autocorrelation is significant at lag 1. The autocorrelation indicates that the model tends to consecutively either overestimate or underestimate the number of defaults. This finding is understandable given that a through-the-cycle (TTC) style PD dampens the sensitivity of the PD to the macroeconomic conditions and does not reflect the improvement or deterioration credit quality, causing a consequent overestimation or underestimation of portfolio losses.

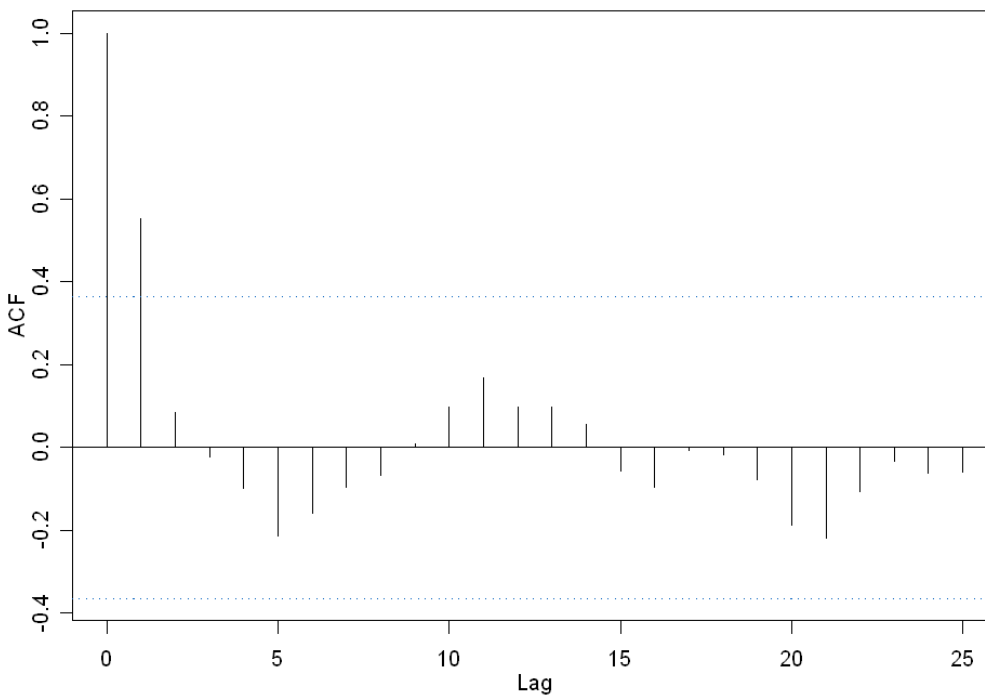


Figure 14 Sample autocorrelation with 95% confidence intervals (Case 1).

Using percentile values, we conduct a two-sided Kolmogorov-Smirnov Goodness-of-Fit test for uniform distribution [0,100] and obtain a p-value of 0.0582. The test statistic falls into the neighboring area of the rejection region if we specify the null hypothesis that the true distribution is a uniform distribution with a significance level of 0.05.

4.2 Point-in-Time Case

We now look at point-in-time (Case 2). In contrast to Figure 13, Figure 15 shows that the percentile points are not serially correlated over time under the Case 2 parameterization. Rather, the percentiles are distributed relatively evenly in the interval of [0, 100], indicating a more accurate assessment of credit risk on the portfolio level. We also see that EDF levels are consistently conservative (i.e. high relative to realized defaults), and the percentile points remain below the 65th percentile of estimated distribution.

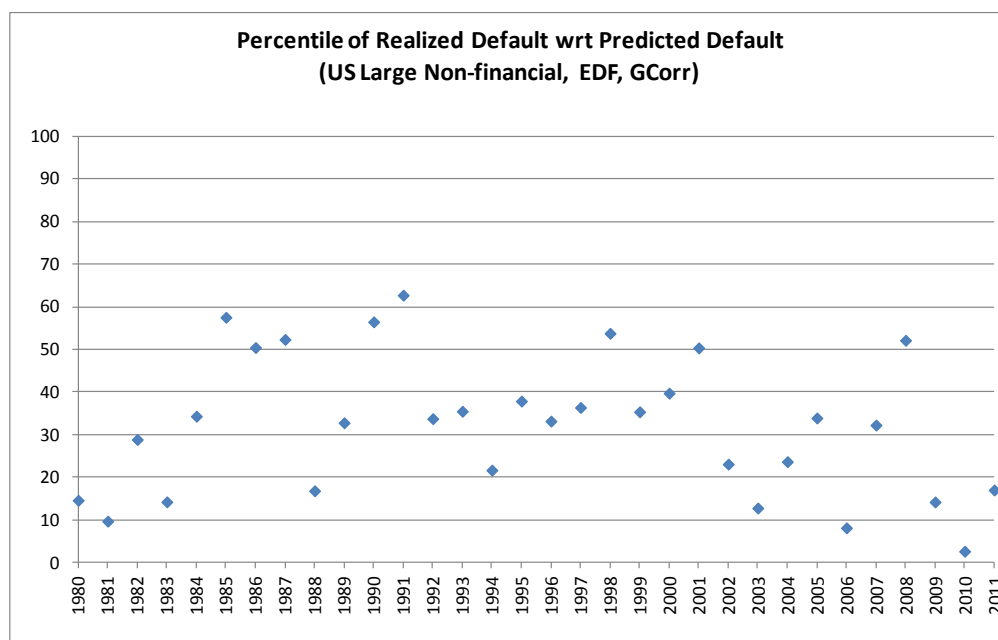


Figure 15 Percentile of realized defaults with respect to predicted default distribution (Case 2).

It has been documented that the EDF credit measure is conservative (i.e., higher than observed default rates even at different annual sales restrictions) due to the hidden defaults issue.¹¹ This failure to capture all defaults can occur for various reasons. For example, when a debt extension occurs, it is difficult for an outsider to know if the extension is caused by the borrower’s inability to pay or by legitimate business need. In other cases, when the loan amount is small, failure to pay is simply written off by the bank, and no public announcement is released. When default data collection relies upon public information to identify defaults, many default events may go missing. This scenario is particularly true for smaller firm borrowers whom draw little public attention.

Moody’s Analytics’ team of specialists aggregates default data utilizing multiple information sources including, but not limited to: bankruptcy newsletters, rating agency debt monitoring publications, news media and news search engines, corporate regulatory filings, internet browsing, and targeted searching. Despite being the largest public default database we are aware of, it is possible that a significant number of defaults are not captured in the data, because in many cases distressed borrowers work out deals privately with lenders, drawing little attention in the media or by data collectors. As there is generally less information available for smaller companies, we believe the hidden default problem is larger for smaller companies. In calibrating the EDF model, we consciously employ only

¹¹ See Crossen, Qu, and Zhang (2011), “Validating the Public EDF Model for North American Corporate Firms.”

large companies in mapping distance-to default to EDF levels in order to circumvent this problem. In addition, the mapping is constructed so the EDF measure is conservative relative to the long-term average default rate, even in the large company sample.

Meanwhile, the EDF credit measures are not too conservative, as we do not observe any extremely low percentiles near the 0th percentile (i.e., over-estimation of the number of defaults) in Figure 15.

Given the non-linear relationship between PD and correlation, we scale the EDF measures to match the average realized default rate during the period 1980–2010, while keeping the GCorr correlation unchanged. We then rerun the portfolios with updated PD and correlation inputs. Figure 16 shows the result. Percentiles are less extreme on the low end of the interval.

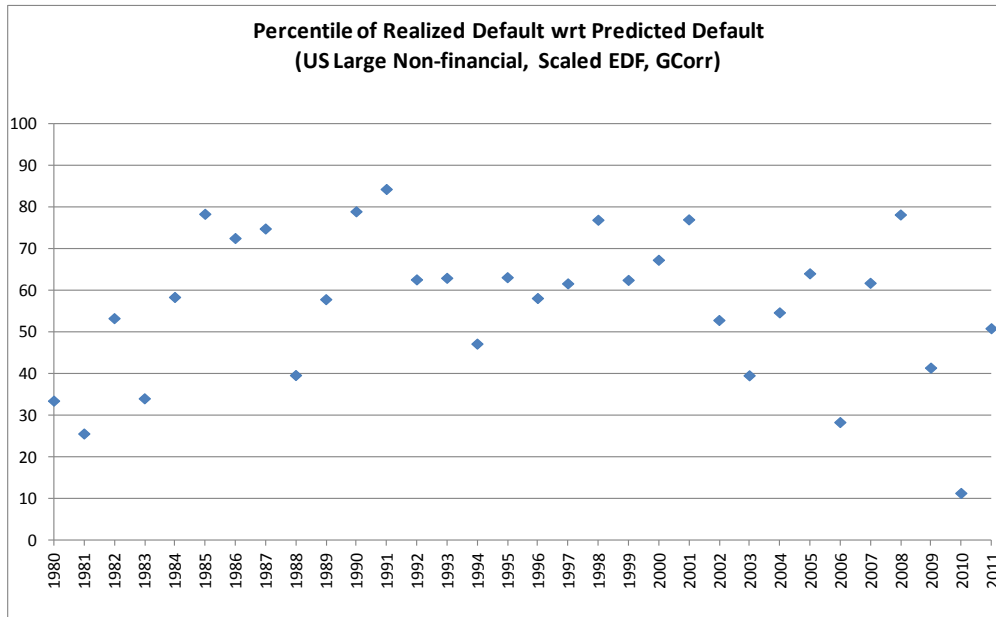


Figure 16 Percentile of realized default with respect to predicted default distribution (Case 2 with scaled EDF).

After matching EDF measures and realized default rate in level, we can focus on examining the effect of correlation upon the distribution of percentile. By visual inspection, we see the moderate level of dispersion among percentiles with the points ranging from the 25th to 85th percentile, and the percentile points are still not serially correlated over time. This randomness is confirmed by computing autocorrelations for percentile values at varying time lags. In Figure 17, none of the autocorrelations at different time lags (lag1 to lag25) significantly differs from zero.

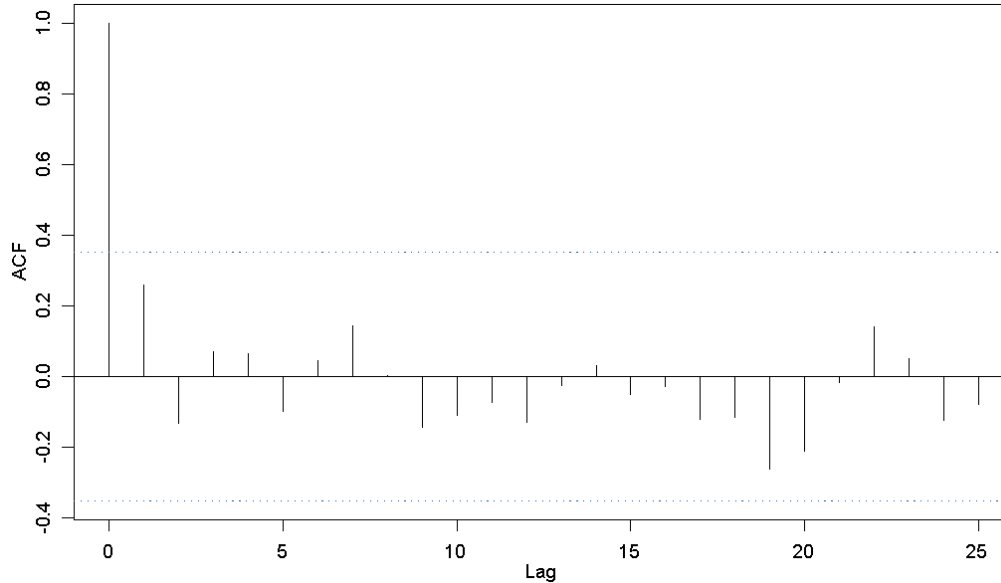


Figure 17 Sample autocorrelation with 95% confidence intervals (Case 2 with Scaled EDF).

Using the percentile values in Figure 16, we also conduct a two-sided Kolmogorov-Smirnov Goodness-of-Fit test for uniform distribution [0,100] and obtain a p-value of 0.047, which rejects the null hypothesis that the true distribution is a uniform distribution with a significance level of 0.05. Given that the clustering of percentile points concentrates around the mean, we explore the uniform range of [10, 90] by repeating the test for uniform distribution [10, 90] and obtain a p-value of 0.233, which fails to reject the null hypothesis at a 5% of significance level.

5 Conclusion

The evaluation of credit portfolio risk model is an important topic. The recent discussion on practices and issues in economic capital frameworks by the Basel Committee highlights this fact. An approach explored by researchers investigates whether or not statistical tests used in market risk can be transferred to evaluate credit risk. The applicability of such an approach for the validation of credit risk models is limited due to the limited number of historical observations available in credit risk.

In this study, we illustrate a validation approach and provide empirical evidence on the ability of credit portfolio model in describing the distribution of portfolio losses. We construct yearly portfolios based upon two typical portfolio parameterizations: (1) a through-the-cycle style parameterization using agency ratings-based long-term average default rates and Basel II correlations and (2) a point-in-time style parameterization using public EDF credit measures and Moody's Analytics Global Correlation Model (GCorr). We then compare the percentiles of realized defaults with respect to the predicted default distributions under different settings.

Results demonstrate that the through-the-cycle style parameterization results in a less conservative view of economic capital and substantial autocorrelation in capital estimates. Results also demonstrate that point-in-time measures help produce consistent and conservative estimates of economic capital over time.

Acknowledgements

The authors thank Christopher Crossen for his comments and suggestions.

Copyright © 2013 Moody's Analytics, Inc. and/or its licensors and affiliates. All rights reserved.

References

- Basel Committee on Banking Supervision (2009). Range of practices and issues in economic capital frameworks.
- Basel Committee on Banking Supervision (2006). International convergence on capital measurement and capital standards.
- Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics* 19, 465-474.
- Crosbie, P., and Bohn, J. (2003). Modeling default risk. Moody's Analytics Technical Document.
- Crossen, C., Qu, S., and Zhang, X. (2011). Validating the public EDF model for North American corporate firms. Moody's Analytics White Paper.
- Crossen, C., and Zhang, X. (2011). Validating the public EDF Model for global financial firms. Moody's Analytics White Paper.
- Frerichs, H., and Löffler, G. (2003). Evaluating credit risk models using loss density forecasts. *Journal of Risk* 5, 4, 1–23.
- Inanoglu, H., and Jacobs, M. (2010). Models for risk aggregation and sensitivity analysis: an application to bank economic capital, *The Journal of Risk and Financial Management*, Vol. 2 (Summer)
- Jacobs, M. (2010). Validation of economic capital models: state of the practice, supervisory expectations and results from a bank study, *Journal of Risk Management in Financial Institutions*, Vol. 3, 4 334–365.
- Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives* 4, 3, 73–84.
- Levy, A. (2008). An overview of modeling credit portfolios. Moody's Analytics White Paper.
- Lopez, J. A., and Saidenberg, M.R. (2000). Evaluating credit risk models. *Journal of Banking & Finance* 24, 151–165.
- Moody's Investor Services, 2011. Corporate default and recovery rates, 1920–2010. Special Comment.
- Moody's Investor Services, 2009. Moody's senior ratings algorithm & estimated senior ratings. Special Comment.
- Rosenblatt, M. (1952). Remarks on a multivariate transform. *The Annals of Mathematical Statistics* 23, 470–472.