

MODELING METHODOLOGY

FROM MOODY'S ANALYTICS
QUANTITATIVE RESEARCH

Authors

Christopher Crossen
Xu 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

Validating the Public EDF Model for Global Financial Firms

Abstract

In this paper, we validate the performance of the Moody's Analytics Public EDF™ (Expected Default Frequency) model for global financial firms during the last decade, including the recent credit crisis and its recovery period. We divide the decade into two sub-periods: an early period, 2001–2007, and a later one, 2008–2010, and then compare the model's performance during these two periods. We focus on the model's ability to prospectively differentiate between defaulters and non-defaulters, its comparison to agency credit ratings, the timeliness of its default prediction, and its accuracy of levels.

Overall, the EDF model's predictive power during the recent sample period has been consistent with its previous longer history, outperforming alternative risk measures, including agency credit ratings. On average, the model provides an early warning signal at least 12 months before default occurs. EDF levels were conservative (i.e., somewhat higher than subsequently realized default rates) before the crisis when compared with later-realized default rates.

We find that EDF credit measures perform consistently well across different time horizons. Our tests indicate that EDF credit measures provide a very useful forward-looking measure of credit risk for global financial firms.

Table of Contents

1	Overview	4
2	EDF Credit Measures' Predictive Power	5
2.1	Data	5
2.2	Validation Results	6
2.3	Rated Firms	8
3	EDF Credit Measures as Early Warning Signals.....	9
4	Validating EDF Level Calibration	10
5	Conclusion	12
	References.....	13

1 Overview

In this paper, we present the results of a study validating the performance of Moody's Analytics' EDF credit measures for global financial firms¹ during the last decade, including the recent credit crisis and subsequent recovery. We look at the period between 2001 and 2010. We divide the decade into two sub-periods: an early period, 2001–2007, and a later one, 2008–2010, and then compare the model's performance during these two time spans. We focus on the model's ability to prospectively differentiate between defaulters and non-defaulters, its comparison with agency credit ratings, the timeliness of its default prediction, and its accuracy of levels.

We validate the EDF model on a regular basis. Our 2007 study examined the period from 2001 through 2006 (Dwyer and Korablev, 2007); our 2009 study focused on the model's performance during the peak of the recent credit crisis (Korablev and Qu, 2009).² This paper details our latest study, which extends our review of the model's performance into the "recovery" period of the credit crisis, including data through 2010. Credit analysis becomes more difficult during high default rate periods, but the severity and complexity of this recent downturn makes default prediction more challenging. In this study, we follow our existing model performance testing methodology, as described in Bohn, Arora, and Korablev (2005) and Dwyer and Korablev (2007).

Overall, we find the EDF model's predictive power during the recent sample period has been consistent with the previous longer history. On average, the model provides an early warning signal at least 12 months before default occurs. EDF levels were conservative (i.e., somewhat high) during the crisis when compared with later-realized default rates.

This paper is organized as follows:

Section 2 presents results for the EDF model's predictive power for global financial firms.

Section 3 discusses EDF credit measures as early warning signals for default.

Section 4 shows simulated level validation results.

Section 5 provides concluding remarks.

¹ "Global financial firms" excludes corporate firms.

² See Korablev and Qu (2009), "Validating the Public EDF Model Performance During the Recent Crisis."

2 EDF Credit Measures' Predictive Power

One of the most important applications of a default prediction model is to prospectively differentiate firms likely to default from those less likely to default. A powerful default prediction model should rank firms from the most risky to the least risky, and this rank ordering should correlate strongly with the subsequent default experience. In this paper, to test the rank order power of the EDF model, we use a well-known approach: the Cumulative Accuracy Profile (CAP). This approach is summarized by a measure known as the Accuracy Ratio (AR). The accuracy ratio ranges between 0 and 1; the closer the AR is to 1, the better the model performs. In extreme cases, for a totally random model that bears no information on impending defaults, $AR=0$. For a perfect model, $AR=100\%$.³

2.1 Data

In this study, to assess the predictive power of the EDF credit measure, we compute Accuracy Ratios on the global financial firms sector. Utilizing rank-ordered EDF measures, we compare Accuracy Ratios for the EDF model in predicting defaults between 2001–2007 and between 2008–2010. All Accuracy Ratios are for a one-year horizon. Our findings show that the inclusion of the crisis period does not reduce the predictive power of the EDF credit measure.

In all tests, we use defaults included in the Moody's Analytics default database, collected and updated daily from numerous printed and online sources worldwide.⁴ As a result, Moody's Analytics employs the most extensive public company default database available. Nevertheless, small public companies often disappear without news or record before they default, or they do not publicly disclose missed payments, which creates a number of hidden defaults, a common challenge often faced by default risk researchers. To reduce this problem of hidden defaults, in many of our tests, we restrict the sample to firms above a certain size threshold, where we believe hidden defaults are less of an issue. For financial firms, we typically use a size threshold of greater than \$100 million or \$300 million in annual sales or book assets, depending upon the context.⁵ Data does not include firms that were "bailed out" by government actions.⁶

In this study, unless otherwise indicated, we calculate Accuracy Ratios by pooling all firm-month observations before ranking them, so that the same EDF value at different time points is assigned the same risk ranking. We track each firm-month EDF observation for 12 months and assign a flag of default or survival. The Accuracy Ratio is computed using the EDF value and default flag pairs, ignoring when the EDF is measured.

The majority of the global financial firms sector consists of U.S. firms. However, we also collect data on firms outside the U.S., including Japan, the United Kingdom, Germany, Canada, Hong Kong, Malaysia, Australia, France, China, Korea, Italy, Switzerland, Taiwan, India, Denmark, Brazil, Thailand, South Africa, and Singapore.

During 2001–2010, there were 430 unique default events (all firm sizes), with 298 during 2001–2007 and 132 during 2008–2010. Table 1 shows the countries and the number of firm-months in each country that constituted the global financial firms module in Moody's Analytics Credit Monitor® and Moody's Analytics CreditEdge® between 2001–2010. As the table shows, outside the U.S., Japan has the largest number observations in the sample, followed by the United Kingdom and Germany.

³ For more details on Accuracy Ratios and a related measure of Receiver Operating Characteristic (ROC), please see Keenan and Sobehart (2000) and Dwyer and Korablev (2007).

⁴ We utilize government filings, government agency sources, company announcements, news services, specialized default news sources, as well as sources within financial institutions to ensure, to the greatest extent possible, that we account for all defaults.

⁵ All figures are in U.S. dollars, unless noted otherwise. We measure size by total annual sales or book assets for Global Financial Firms non-financial firms. Where the firm's total sales number is not available, we use book assets.

⁶ For results that include bailouts, please refer to Korablev and Qu (2009).

Table 1 Countries in the Global Financial Firms Database: 2001–2010

Country	Number of Observations (firm-months)	Country	Number of Observations (firm-months)
United States	108,378	Korea	6,898
Japan	32,059	Italy	6,182
United Kingdom	17,309	Switzerland	5,955
Germany	9,782	Taiwan	5,915
Canada	9,499	India	5,679
Hong Kong	9,471	Denmark	4,587
Malaysia	9,222	Brazil	4,570
Australia	8,678	Thailand	4,471
France	8,322	South Africa	4,424
China	7,403	Singapore	3,879

2.2 Validation Results

We track the solvency of each firm's EDF credit measure for 12 months. For the 2001–2007 measurement period, we use EDF credit measures between December 2000 and December 2006 and default events between January 2001 and December 2007. Similarly, for the 2008–2010 period, we use EDF credit measures between December 2007 and December 2009 and default events between January 2008 and December 2010. Although comparing Accuracy Ratio statistics between two times periods is problematic, it provides a sense of the model's relative performance across time.

In Figure 1, the left panel shows the EDF model's performance during 2001–2007. The right panel presents model performance during 2008–2010. Both samples include only firms with greater than \$100 million in book assets. During the later, crisis period, the EDF credit measure's Accuracy Ratio is 74.3%, compared with the 82.5% Accuracy Ratio during the 2001–2007 period. For the second sample period, the dramatic change in credit conditions for financial companies, beginning with the unprecedented benign period of 2007 into the largest financial crisis post the Great Depression, has challenged the EDF model. Although problems may arise when comparing Accuracy Ratios from two different data samples, comparing the recent sample, which includes the recent credit crisis, with the prior sample, shows that the EDF model performs as expected during both periods.⁷

⁷ We choose the version of Altman's Z-Score designed for public firms, which includes market capitalization in the leverage ratio. For the specific form we use, see Dwyer and Korablev (2007).

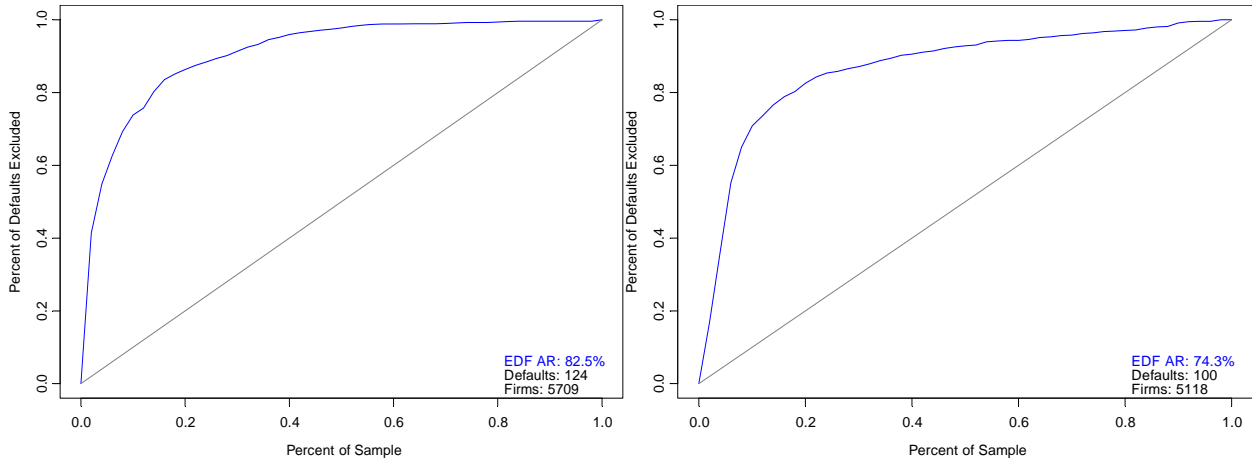


Figure 1 CAP Curves Comparing EDF Credit Measure Historical Performance for Global Financial Firms, Book Assets >\$100 million: 2001–2007 (left panel) versus 2008–2010 (right panel).

Figure 2 shows the CAP Curves of the historical performance of global financial firms, book assets >\$100 million, 2001–2010.⁸

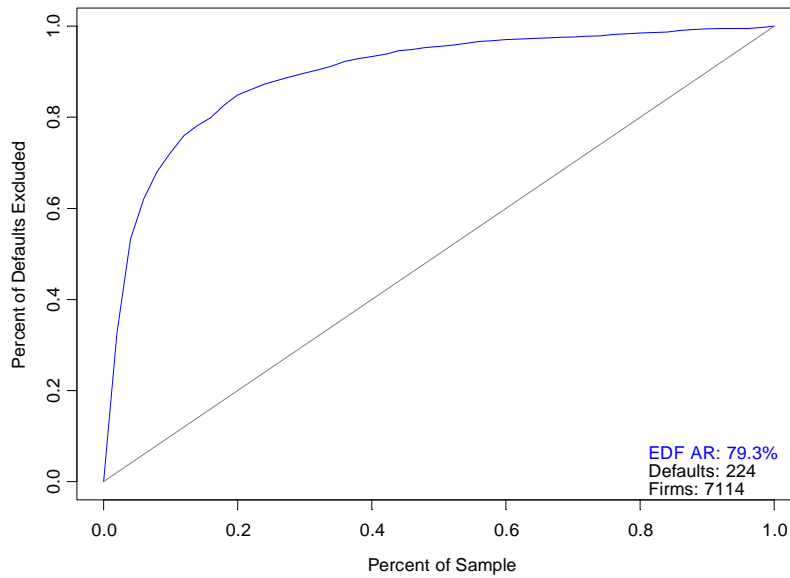


Figure 2 CAP Curve Comparing EDF Credit Measure Historical Performance for Global Financial Firms, Book Assets >\$100 million: 2001–2010.

⁸ As a result of the entry and exit of firms over time, the sample of firms used for the validation changes to some extent over time as well.

Table 2 shows the Accuracy Ratios for the EDF credit measure at the end of each year between 2000 and 2009.

Table 2 Accuracy Ratios for EDF credit measures by year, global financial firms, book assets > \$100 million: 2001–2009

Date (end of)	Numbers of Companies	Numbers of Defaults	Default Rates	Accuracy Ratios EDF
2000/12	3,173	30	0.95%	75.5%
2001/12	3,324	32	0.96%	83.0%
2002/12	3,351	23	0.69%	74.1%
2003/12	3,386	7	0.21%	92.8%
2004/12	3,541	4	0.11%	79.2%
2005/12	3,676	6	0.16%	96.4%
2006/12	3,995	9	0.23%	74.3%
2007/12	4,289	36	0.84%	68.2%
2008/12	4,375	36	0.82%	82.8%
2009/12	4,291	25	0.58%	86.0%

2.3 Rated Firms

In this section, we focus on the subset of firms with both EDF credit measures and Moody's Investors Service credit ratings. Figure 3 presents the CAP Curve for rated global financial firms, 2001–2010, derived from 34 defaults and 697 unique firms.⁹

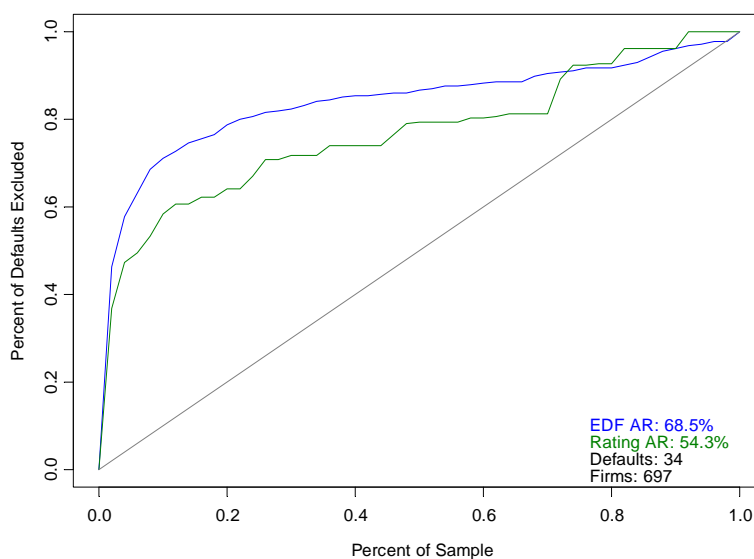


Figure 3 CAP Curves Comparing EDF Credit Measure Historical Performance for Global Financial Firms Rated by Moody's Investor Services: 2001–2010.

⁹ Ratings may be withdrawn before a default event occurs. We count a firm as a rated default if it was rated within 12 months prior to its default.

3 EDF Credit Measures as Early Warning Signals

To test the timeliness of EDF credit measures as early warning signals, we create a sample of defaulted firms from January 2008 through December 2010. We compute the 25th, 50th, and 75th percentiles of the EDF credit measure for these defaulted firms dating back to December 2001. We also compute the same percentiles for the entire sample, and then plot two sets of percentiles on the same graph. If the EDF credit measure provides early warning signals, we expect the distribution of EDF credit measures for the defaulted firms to (1) be higher relative to the distribution for all firms, and (2) react sooner to adverse changes in credit risk than the entire population as they approach default dates. These findings are indeed what we observe.

In the global financial firms sector, 132 firms defaulted between January 2008 and December 2010. Figure 4 presents EDF credit measure percentiles for these defaulted firms. The red lines represent the 25th, 50th, and 75th percentiles of EDF credit measures for companies that defaulted between January 2008 and December 2010. The blue lines represent percentiles for the entire sector.

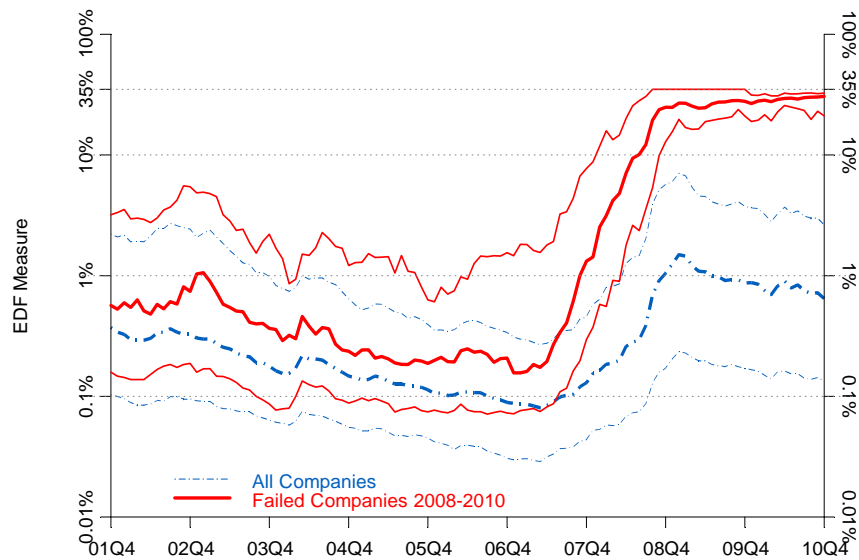


Figure 4 Early Warning and Distribution of EDF Credit Measures for Global Financial Firms: 25th, 50th, and 75th percentiles.

These two distributions are distinctly different. As shown in Figure 4, beginning in mid-2004, defaulters were riskier than the rest of the sample and usually had higher EDF credit measures than non-defaulters, well before default. As the entire sector improved between the end of 2002 and mid-2007, defaulters actually began to deteriorate during mid-2006. The top 25th percentile of the defaulters' EDF measures began increasing as far back as late-2005. The speed of the deterioration increased dramatically in mid-2007. Defaults were realized between January 2008 and December 2010. The entire sector's risk began to increase beginning in mid-2007 and remains at high levels.

Another way to test EDF credit measures' early warning power is to construct an event study for default firms around the date of default. In this test, we create a sample of defaulted firms, retaining monthly observations from 24 months prior to default, and then compute the median EDF credit measure by month up to the time of default ($t = 0$). Figure 5 overlays and compares the median EDF credit measures for defaulted firms for the 2001–2007 and 2008–2010 time periods.

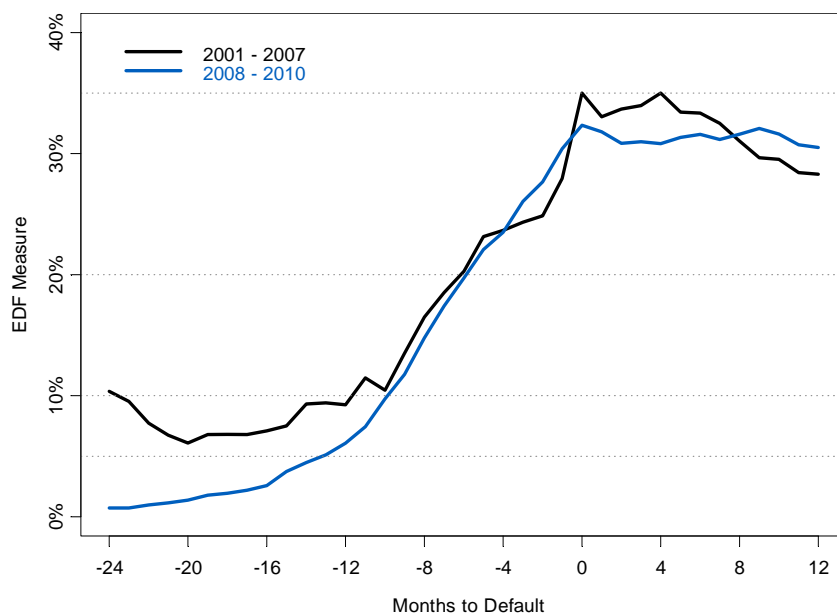


Figure 5 Median EDF Credit Measures for Global Financial Firms Defaulter

Figure 5 demonstrates that, in the event of default, EDF credit measures provide early warning signals for the 2001-2007 period, as well as the most recent crisis period. For both time frames, the EDF credit measure is elevated more than 12 months prior to the credit event and continues increasing steadily. Figure 5 shows that the median EDF level 24 months prior to default is lower for the 2008–2010 period when compared with the 2001–2007 period. This difference is not surprising, given that the two years prior to the crisis, 2005–2007, were a relatively benign period. Additionally, in 2008–2010, the slope of the increase in median EDF level steepens as early as 16 months before default occurs.

4 Validating EDF Level Calibration

In this section we test the public firm EDF model's level calibration. EDF credit measures are forward-looking probability measures. Thus, for a given portfolio of companies, the model implies a distribution of possible default rates. Here we test whether the realized default rate is statistically consistent with such a default rate distribution implied by the EDF credit measures using simulated levels.

For each year in the sample, we take EDF values calculated at the beginning of the year, use them to simulate a distribution of default rates and then compare the actual default rate during the year against the resulting distribution. We use a pair-wise asset correlation of 0.19, calibrated using long-term data from the Moody's Analytics Global Correlation Model (GCorr). For each year, we run 1,000 simulations to create a distribution of 1,000 simulated default rates. We can use a simple numerical example to illustrate the nature of the test. Suppose we have 100 independent companies, each with a default probability of 10%. The expected default rate is 10%, and it is possible, but not likely, to have a default rate of less than 5% (the p-value is only 5.75%).¹⁰ When we actually observe a 5% default rate, we would reject, at a 10% confidence interval, the hypothesis that the default probabilities are 10%.

When companies are correlated, the implied default rate distribution is wider, in that, the likelihood is higher for the realized default rate to deviate from the mean prediction. For example, for two independent firms, each with a 50% true default probability, the likelihood of both (or neither defaulting) defaulting is only 25%. However, if these two firms are

¹⁰ The likelihood of less than n% default is $\sum_{i=0}^n \frac{100!}{i!(100-i)!} (0.1)^i (0.9)^{100-i}$.

perfectly correlated, the likelihood of both defaulting (or neither defaulting) increases to 50%. Therefore, the likelihood of deviation from the mean prediction increases with correlation.

Based on the above concepts, Kurbat and Korablev (2002) developed a method that uses realized defaults for testing default probability models. Specifically, they assume firm asset values are correlated with one common factor, and obtain a default rate distribution by simulating random realizations of a common factor and firm-specific factors. In simulations, each firm’s unconditional default probability is kept at the level predicted by the subject probability of default (PD) model. In this study, we use the same approach.

Figure 6 presents results. To help address the hidden defaults issue, we limit our sample to companies with book assets greater than \$300 million. As shown in Figure 6, the realized default rate was mostly lower than the median EDF level during the past ten years, and it remained within the intervals bounded by the 10th and 90th percentiles of the simulations, with the exception of 2004–2005 and 2009-2010. In general, EDF levels tend to be higher than observed default rates, especially preceding market recoveries.

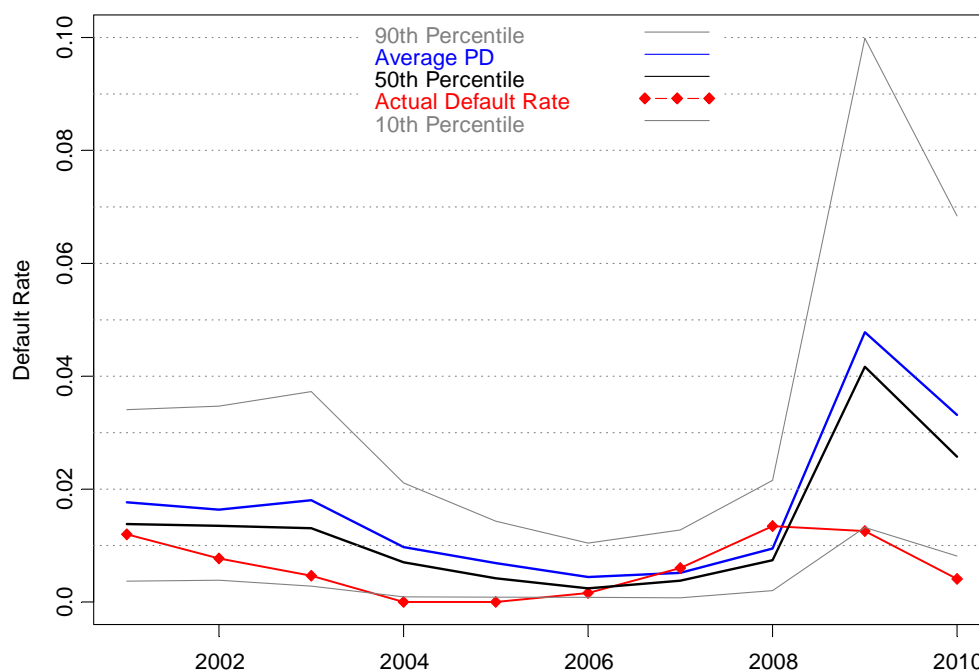


Figure 6 Default Rate for Global Financial Firms, Book Assets > \$300 million, 2001–2010.

We see that EDF levels are higher than observed default rates. This finding may be partially due to the hidden defaults issue. Hidden defaults refers to the failure of a default dataset to capture all economic defaults. This failure 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 on public information to identify defaults, many default events may go missing. This is particularly true for smaller firm borrowers that draw little public attention.¹¹

¹¹ See Stein and Dwyer (2005) and Dwyer and Qu (2007) for more information regarding hidden defaults.

5 Conclusion

In this study, we test the public EDF model for global financial firms using major performance measures: Accuracy Ratios in default prediction, early warning signals, comparison with agency ratings, and default risk levels, with special focus on the most recent credit crisis and recovery period.

In Accuracy Ratio testing, we find that EDF credit measures are as powerful as they have been historically in their ability to prospectively differentiate between defaulters and non-defaulters. We find that the EDF model provides early warning signals. The distribution of EDF levels for defaulters begins to emerge out of the entire population distribution more than 12 months before defaults occur. EDF levels were conservative (i.e., somewhat high) before the crisis, compared with later realized default rates.

We find that the EDF levels are consistently higher than observed default rates, due to downwardly-biased default rate observations, caused by the hidden defaults issue, as well as conservatism built into the EDF model. Over the longer history, the realized default rate, with better default coverage, typically lies within the prediction interval, and we cannot reject the hypothesis that the EDF values are true measures of default risk.

Acknowledgements

We are grateful to everyone who contributed to this paper. All remaining errors are, of course, our own.

Copyright © 2011 Moody's Analytics, Inc. and/or its licensors and affiliates. All rights reserved.

References

- Arora, Navneet, Jeffery Bohn, and Irina Korablev, "Power and Level Validation of the EDFTM Credit Measure in the U.S. Market." Moody's KMV, 2005a.
- Crosbie, Peter and Jeffrey Bohn, "Modeling Default Risk." Moody's KMV, Revised December 2003.
- Dwyer, Douglas and Irina Korablev, "Power and Level Validation of Moody's KMV EDFTM Credit Measures in North America, Europe, and Asia." Moody's KMV, September 2007.
- Dwyer, Douglas and Shisheng Qu, "EDFTM 8.0 Model Enhancements." Moody's KMV, January 2007.
- Keenan, Sean and Jorge Sobehart, "A Credit Risk Catwalk." Risk, July 2000.
- Korablev, Irina and Shisheng Qu, "Validating the Public EDF Model Performance During the Recent Credit Crisis," June 2009.
- Kurbat, Matt and Irina Korablev, "Methodology for Testing the Level of the EDF Credit Measure." Moody's KMV, August 2002.
- Stein, Roger and Douglas Dwyer, "Inferring the Default Rate in a Population by Comparing Two Incomplete Default Databases." Moody's KMV, October 2005.