OCTOBER 2011

MOODY'S ANALYTICS

MODELING METHODOLOGY

FROM MOODY'S ANALYTICS QUANTITATIVE RESEARCH

Validating the Public EDF Model for European Corporate Firms

Authors

Christopher Crossen

Xu Zhang

Contact Us

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

Europe
+44.20.7772.5454
clientservices.emea@moodys.com

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

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

Abstract

In this paper, we validate the performance of the Moody's Analytics Public EDF™ (Expected Default Frequency) model for European corporate 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, 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. On average, the model provides an effective early warning signal beginning 12 months before default occurs. EDF levels were conservative (higher than subsequently realized default rates) during the crisis when compared with 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 firms in Europe.

Table of Contents

1 Overview

In this paper, we present the results of a study validating the performance of Moody's Analytics' EDF credit measures for European corporate firms[1] during the last decade, including the recent credit crisis and subsequent recovery. We look at the period between 2000 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, 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] and Gokbayrak and Chua, 2009).[3] 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 particularly 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 European corporate 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] "European corporate firms" excludes financial firms.

[2] See Korablev and Qu (2009).

[3] See Gokraybak and Chua (2009).

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%.[4]

2.1 Data

In this study, to assess the predictive power of the EDF credit measure, we compute Accuracy Ratios on the European corporate 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.[5] 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. We typically use a size threshold of greater than $100 million or $300 million in annual sales, depending upon the context.[6]

In this study, we calculate Accuracy Ratios by pooling all firm-month EDF measure observations. We first rank the EDF credit measures for each year-month, attain the percentile, and then pool the credit measures together to conduct the Accuracy Ratio test. The same EDF values at different points in time most likely have different percentiles, so their relative risk ranking will be different. Given a default, we flag all the EDF observations up to 12 months prior to the default date and discard any post-default observations within two years. We then compute the AR using the EDF ranking percentiles and default flag pairs, ignoring when the EDF measure is recorded.

Most of the European corporate sector consists of firms from the United Kingdom. However, we also collect data on firms outside the U.K. including Germany, France, Italy, Greece, Sweden, South Africa, Switzerland, Israel, Russian Federation, Turkey, Netherlands, Norway, Poland, Finland, Spain, Denmark, Belgium, Austria, and Portugal.

During 2001–2010, there were 606 unique default events (all firm sizes), with 519 during 2001–2007 and 87 during 2008–2010. Table 1 shows the countries and the number of firm-months in each country that constituted the European module in Moody's Analytics Credit Monitor® and Moody's Analytics CreditEdge® between 2001–2010. As the table shows, outside the U.K., Germany has the largest number observations in the sample, followed by Germany, France, and Italy.

---

[4] 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).

[5] 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.

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

Table 1 Countries with the Most Observations in the European Database: 2001–2010

Country	Number of Observations (firm-months)	Country	Number of Observations (firm-months)
United Kingdom	88,380	Turkey	14,277
Germany	53,847	Netherlands	13,835
France	53,127	Norway	12,358
Italy	21,038	Poland	11,799
Greece	20,305	Finland	11,236
Sweden	19,743	Spain	11,033
South Africa	19,336	Denmark	9,786
Switzerland	17,706	Belgium	9,085
Israel	15,083	Austria	6,803
Russia	14,370	Portugal	4,922

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.

There are different approaches to measuring Accuracy Ratios over time. One can pool all the data, and then one is testing the ability of the model to rank order credits both cross-sectionally and over time. A second method is to first rank companies by their EDF values within a time period, and then pool the data and compute the accuracy ratio using the ranking rather than the actual EDF. A third method is to compute the accuracy ratio at each point in time and take the average over time. The latter two approaches isolate the model's ability to rank order risk in a cross sectional context. For purposes of this paper, we adopt the second approach. We address the time series properties of the model in the context of level validation in Section 4.

Figure 1, 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 annual sales. During the later, crisis period, the EDF credit measure's Accuracy Ratio is 67.2%, compared with the 82.3% Accuracy Ratio during the 2001–2007 period. 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 better than the Z-Scores 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).

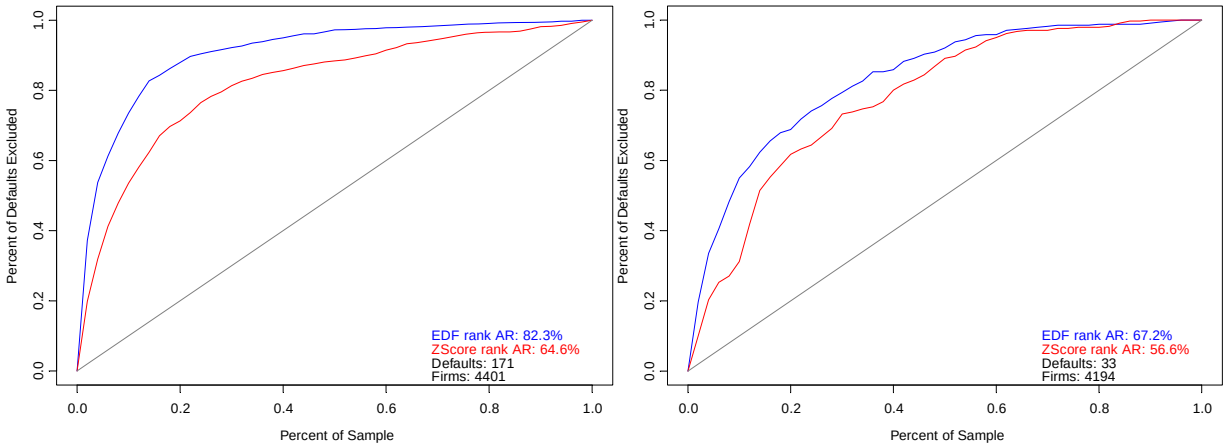

Figure 1 CAP Curves Comparing Rank-Order Power EDF Credit Measures and Z-Scores of European Corporate Firms, Annual Sales>$100 million: 2001–2007 (left panel) versus 2008–2010 (right panel).

Figure 2 shows the CAP plot for rank-ordered EDF credit measures and Z-Scores, European corporate firms, annual sales>$100 million, 2001–2010. The EDF credit measure's Accuracy Ratio is 78.8%, versus the 62.9% Z-Score. Over the entire period, EDF credit measures outperform Z-Scores. The results suggest that the EDF model performs consistently over time and in different credit cycles.[8]

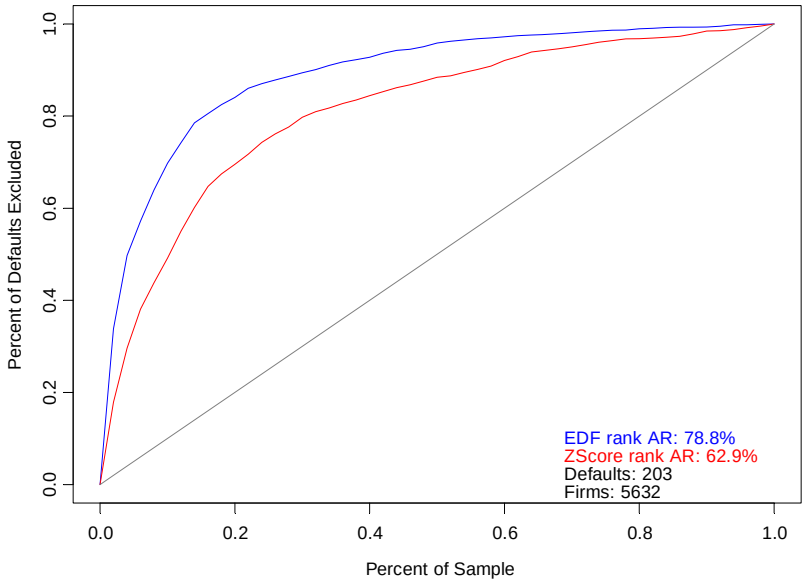

Figure 2 CAP Plot for Rank-Ordered EDF Credit Measures and Z-Scores, European Corporate Firms, Annual Sales>$100 million: 2001–2010.

[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 and the Z-Score at the end of each year between 2000 and 2009, annual sales >$100 million. We can see that the EDF credit measure outperforms the Z-Score during all years.

Table 2 Accuracy Ratios for EDF credit measures and Z-Scores by year, European corporate firms, annual sales>$100 million: 2000–2009

Date (end of)	Numbers of Companies	Numbers of Defaults	Default Rates	Accuracy Ratios EDF	Accuracy Ratios Z-Score	Accuracy Ratio Differences
2000/12	2,452	32	1.3%	76.5%	43.7%	32.8%
2001/12	2,545	40	1.6%	86.3%	64.9%	21.3%
2002/12	2,687	23	0.9%	74.6%	70.6%	4.0%
2003/12	2,756	30	1.1%	73.6%	69.9%	3.7%
2004/12	2,843	11	0.4%	93.6%	83.6%	10.0%
2005/12	2,785	9	0.3%	90.1%	61.6%	28.4%
2006/12	3,086	3	0.1%	34.3%	14.9%	19.4%
2007/12	3,459	6	0.2%	71.9%	67.3%	4.5%
2008/12	3,477	22	0.6%	73.2%	59.6%	13.6%
2009/12	3,480	2	0.1%	89.1%	87.8%	1.3%

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 European corporate sector, 87 firms defaulted between January 2008 and December 2010. Figure 3 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.

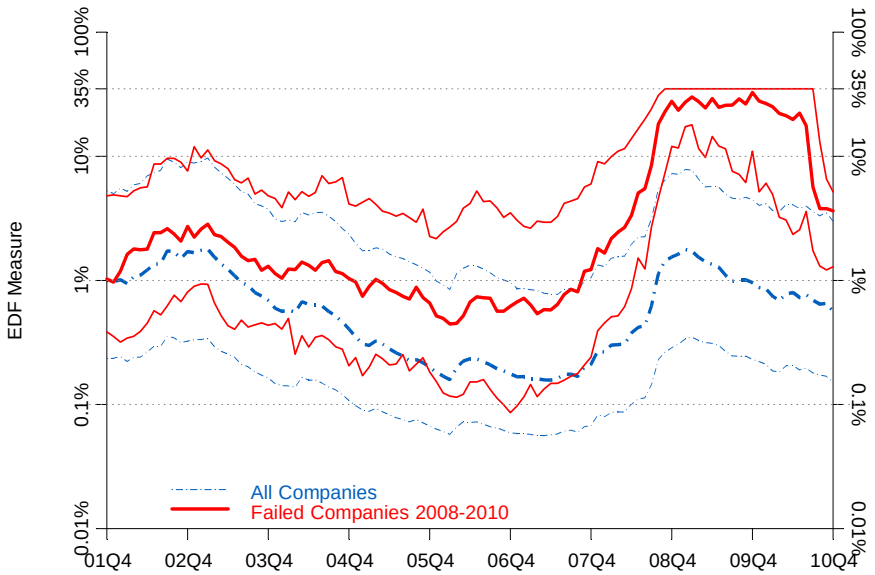

Figure 3 Distribution of EDF Credit Measures for European Corporate Firms: 25th, 50th, and 75th percentiles.

These two distributions are distinctly different. As shown in Figure 3, 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 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 began to decrease beginning in early 2009.

In Figure 3, we also observe a steep drop in the 25th percentile of the defaulter's EDF credit measures, beginning in late 2009. This finding is mainly due to some surviving companies emerging from bankruptcy that were then relisted in the equity market, and thus, their EDF levels improved significantly. For example, IQ Power AG's EDF credit measure changed from 34.5% in December 2009 to 1.0% in October 2010, a result of its reorganization and recapitalization.

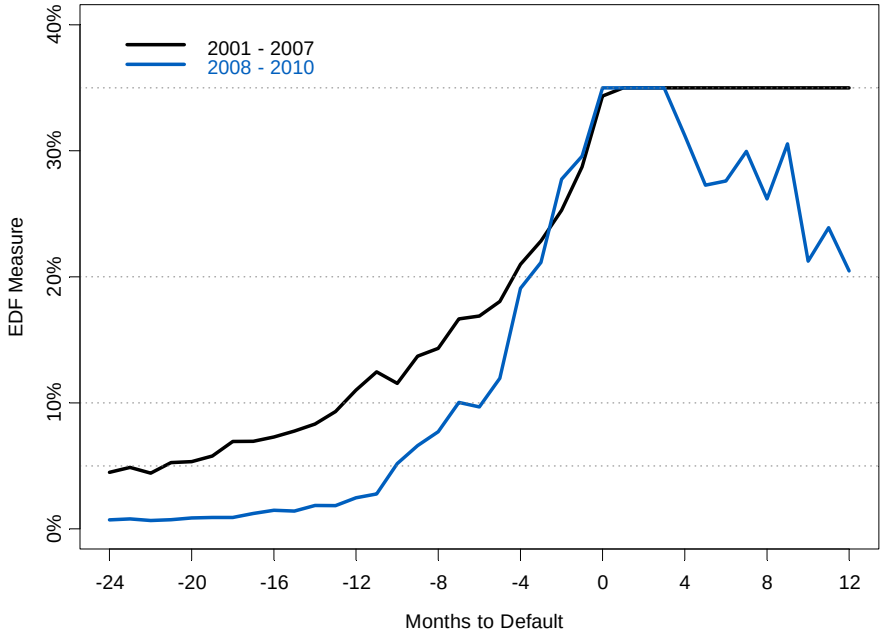

Figure 4 Median EDF Credit Measures for European Corporate Defaulters.

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 4, overlays and compares the median EDF credit measures for defaulted firms for the 2001–2007 and 2008–2010 time periods.

Figure 4 demonstrates that, in the event of default, the early warning performance of EDF credit measures, including the recent recession, is similar to historical performance. For both periods, the EDF credit measure is elevated more than 12 months prior to the credit event and continues increasing steadily. Figure 4 shows that the median EDF level 24 months prior to default is lower for the 2008–2010 period 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 11 months before default occurs. Comparing the two sample periods, recent defaulters were much safer two years before defaults and they experienced sharper increases in EDF measures.

The rapid drop of median EDF for 2008–2010 defaulters was caused primarily by a few companies whom emerged from bankruptcy and survived the sampling, and whose EDF levels dropped significantly after restructuring, from more than 20% to a few percent or even lower.

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%).[9] 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 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 5 presents results. To help address the hidden defaults issue, we limit our sample to companies with annual sales greater than $300 million. As shown in Figure 5, 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 2003 and 2009–2010. In general, EDF levels tend to be higher than observed default rates.

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

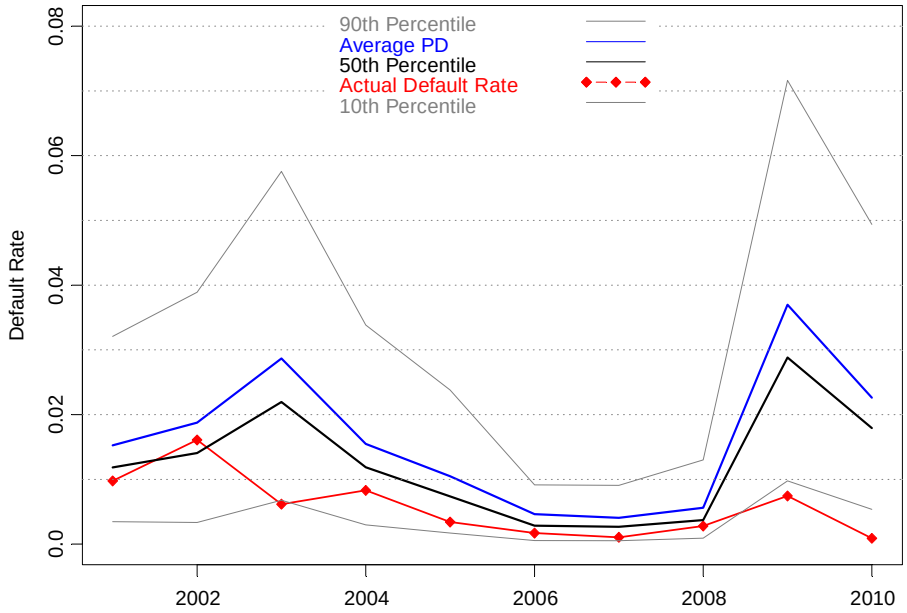

Figure 5 Default Rate for European Corporate Firms, Annual Sales>$300 million, 2001–2010.

We see that EDF levels are higher than observed default rates (even at different annual sales restrictions). 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.[10]

5 Conclusion

In this study, we test the public EDF model for European corporate firms using three major performance measures: Accuracy Ratios to test rank order power, early warning signals, 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.

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

Acknowledgements

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

References

Arora, Navneet, Jeffery Bohn, and Irina Korablev, "Power and Level Validation of the EDF™ 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 EDF™ Credit Measures in North America, Europe, and Europe." Moody's KMV, September 2007.

Dwyer, Douglas and Shisheng Qu, "EDF™ 8.0 Model Enhancements." Moody's KMV, January 2007.

Gokbayrak, Ozge, Lee Chua, "Validating the Public EDF™ Model During the Credit Crisis in Asia and Europe." Moody's Analytics White Paper, 2009.

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.