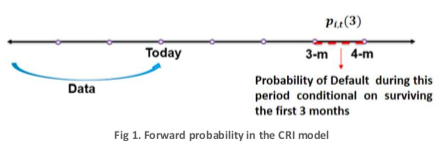
Probability of Default

1. Abstract: Probability of Default (PD) is the core credit product of the Credit Research Initiative(CRI). The CRI system is built on the forward intensity model developed by Duan(2012. Journal of Econometrics). This white paper describes the foundations and numerical realization are presented in RMI-CRI Technical Report (Version 2017). This white paper contains 5 sections. Among them, section 2&3 describes the methodology and performance of the model respectively, and section four relates to examples of how CRI PD can be used.
2. Overview: Probability of Default is the core credit measure of CRI corporate default prediction system built on the forward intensity model of Duan. This forward intensity model is governed by two independent doubly stochastic [Poisson processes](http://www.ruanyifeng.com/blog/2015/06/poisson-distribution.html), operating on forward time instead of spot time. This enables the model to produce forward-looking PD-term structures of public firms based on dynamic learning from the macrofinancial and firm-specific data.

* Key features of the CRI model are:
* Combines a reduced-form model (based on a forward intensity construction) and a [structural model](https://books.google.com.sg/books?id=cxew92azWWMC&pg=PA369&lpg=PA369&dq=简化式模型&source=bl&ots=MW1tHfLJ7Q&sig=ACfU3U0spcNhvoerAPD73NaI1ahwLkPZiw&hl=zh-CN&sa=X&ved=2ahUKEwj89bKVk-bkAhVR73MBHeIkBesQ6AEwAXoECAkQAQ%20-%20v=onepage&q=简化式模型&f=false) (using Distance-To-Default as one of its input covariates)
* Accommodates the two risks that a listed firm might encounter; namely default risk and risk of other types of corporate exists (i.e. mergers and acquisitions).
* Uses forward probabilities of default and other types of exists as building blocks to construct the PD term structure in a consistent manner.
* Employs sixteen input covariates (or default predictors) from both market-based and accounting-based firm-specific attributes, as well as macro-financial factors (except for China which have fourteen covariates)
* Coverage:
* July 2010, 17,000 public firms among 12 Asian economies.
* Jan 2018, 67,000 exchange-listed firms in 128 economies worldwide, 1 month—5years
* Out of those firms, over 34,000 are currently actively listed and have their PD updated on a daily basis.

1. Methodology:

* Model description: the building block of the CRI corporate default prediction model is the conditional forward probability. As figure 1 illustrates, when firm i is at time t facing the future, pi,t(3) is the probability that the firm defaults in the fourth month, conditional on its survival up to the third month.



Formally, for each forward period T, pi,t(T) is constructed on a forward intensity function, whose inputs include the state of economy (macrofinancial risk factors Xt,) and the vulnerability of individual obligors (firm-specific attributes Yi,t):

Pi,t(T) = PT(Xt, Yi,t)

With Pi,t(T) in place, the multi-period default probabilities with different term structures can be obtained through the typical survival-exit formula. The underlying forward intensity functions are parameterized, and the parameters are estimated for each calibration groups on a monthly basis as new information accumulated in the CRI database.

* Input Covariates

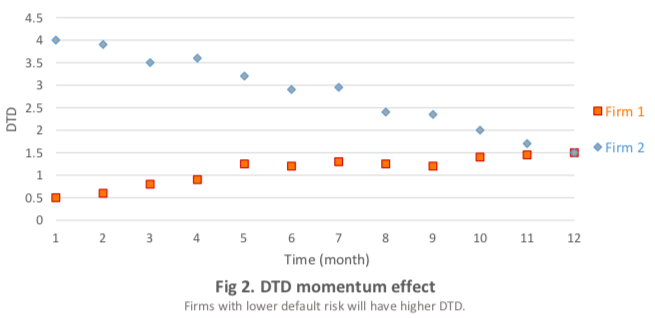
Following the notation above, firm I’s input covariates at time t are represented by 1) the vector Xt that is common to all firms in the same economy2, and 2) a firm-specific vector Yt with components constructed from the firm’s financial statements and market capitalizations. The CRI corporate default prediction model employs four macrofinancial variables and twelve firm-specific variables, as presented in Table 1 below.

|  |  |  |
| --- | --- | --- |
| **Table 1. Input covariates for the CRI PD model** | | |
|  | page5image1981023248  **Model Inputs** | **Description** |
| **Macro- Financial Factors**  page5image1981027056 | page5image1980887136  Stock Index Return？ | Trailing 1-year return of the prime stock market, winsorization and currency adjusted |
| page5image1980990864  Short-term Risk-Free Rate  page5image1981034640 | Yield on 3-month government bills |
| Economy-level Distance-To-Default for financial firms | Median Distance-to-Default of financial/non- financial firms in each economy/country inclusive of those foreign firms whose primary stock exchange is in this economy/country (Not applicable to China) |
| Economy-level Distance-To-Default for non-financial firms |
| page5image1980022704page5image1980020816  **Firm-Specific Attributes**  page5image1980026672page5image1980027440page5image1980027824page5image1980028336 | page5image1980030752  Distance-to-Default (level)  page5image1980035584 | Volatility-adjusted leverage based on Merton (1974) with **special** treatments |
| Distance-to-Default (trend) |
| Cash/Total Assets (level) | For financial firm’s liquidity - Logarithm of the ratio of each firm’s sum of cash and short-term investments to total assets |
| page5image1981219328  Cash/Total Assets (trend)  page5image1981222704 |
| Current Assets/Current Liabilities (level) | For non-financial firm’s liquidity - Logarithm of the ratio of each firm’s current assets to current liabilities |
| Current Assets/Current Liabilities (trend) |
| Net Income/Total Assets (level) | Profitability - Ratio of each firm’s net income to total assets |
| Net Income/Total Assets (trend) |
| Relative Size (level) | Logarithm of the ratio of each firm’s market capitalization to the economy’s median market capitalization over the past one year |
| Relative Size (trend) |
| Relative Market-to-Book Ratio | Individual firm’s market misvaluation/ future growth opportunities relative to the economy’s median level of market-to-book ratio |
| Idiosyncratic Volatility | 1-year idiosyncratic volatility of each firm, computed as the standard deviation of its residuals using the market model |

In the table above, “level” is computed as the 12-month moving average (a minimum of six observations in the 12-month range are required, otherwise level variables will bear missing values), and “trend” is computed as the current value minus the “level” value (if the current month value is missing, the trend variable is set to be the last value in the previous month).

The “trend” measure captures the momentum effect and gives a hint about the direction of future movements. Duan at aL(2012) show that using the level and trend of the measure for some input covariates significantly improves the predictive power of the model, particularly for short horizons.

In order to understand the momentum effect, consider the case of two firms that have the same current value of Distance-To-Default (DTD). Firm 1 reaches its current value of DTD from a lower level, while Firm 2 reaches the same current value of DTD as Firm 1 but from a higher level, as shown in Figure 2. If only the current value of the DTD is employed for default prediction, the impact of the DTD on PD would be identical for both firms. Intuitively however, one would expect the DTD of Firm 1 to have an upward momentum whereas that of Firm 2 to continue its decline. In order to account for such momentum effects, the CRI uses both level and trend attributes in the computation of its PD.



DTD has long been recognized as an important indicator of a firm’s credit quality and is employed by the CRI as a default predictor in the forward intensity model.

Typically, the DTD for each firm is estimated using Merton’s structural model with the same assumptions on the debt maturity and size in the KMV implementation, i.e.,

DTDt =

*Where Vt is the asset value following a geometric Brownian motion with drift and volatility , L is the default point with value equal to short-term liabilities plus half of long-term liabilities, and is set to 1 year.*

However, to improve the traditional DTD measure, the CRI implements some special treatments on its own DTD calculation to overcome several drawbacks that have been identified in the literature.

The key treatments are:

* Follow Duan (2010) to add a fraction () of other liability to the KMV default point L
* Set = to improve the stability of DTD estimation
* Standardize the firm’s market value by its book value to handle the scale change due to any major investment and financing action

The parameters required for DTD estimation are estimated by the maximum likelihood method described in Duan (1994, 2000)

A brief expression of of the CRI’s version of DTD can be rewritten as:

DTD =

*Where the default point is set to L = Current Liabilities+1/2 Long-term Liabilities + \* Other Liabilities, and is shared by firms in the same sector (financial vs. non-financial) in each calibration froup.*

* Calibration

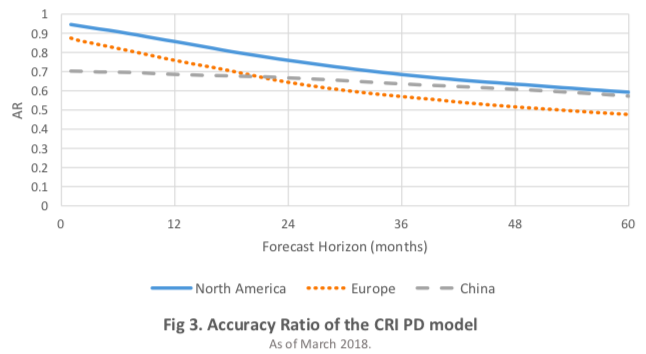
Currently, the data for the CRI corporate default prediction system comes from various international data distributors. it is worth noting that there are few or no credit default events in certain economies due to limited number of listed firms in those counties, which means that calibrating models for individual economy would not be statistically meaningful. In view of this, public companies around the world are segregated into six calibration groups according to certain similarities in the stage of economic development and geographic locations of their primary exchanges. These calibration groups are North America, Europe, Asia-developed economies, Emerging Markets, China, and India.

The CRI PD measures of companies within the same calibration group share the same set of parameters, (except for some covariates in some special circumstances). To overcome optimization difficulties caused by high dimensionality of the parameters (i.e. 16 covariates in general and 14 covariates in the case of China for 60 monthly prediction horizons), the CRI system employs the Nielson-Siegel term structure function and relies on sequential Monte Carlo for the model’s estimation.

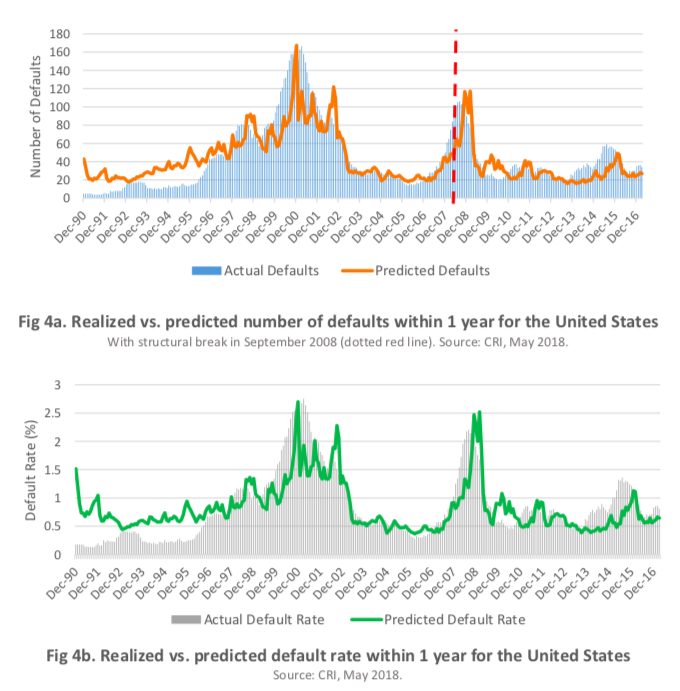
1. Model Performance

The Accuracy Ratio (AR) is a popular quantitative measure for evaluating the discriminatory power of a default prediction system. It is the ratio of (a) the differential of the performance of the evaluated system and the random system over (b)the differential of the performance of the perfect system and the random system. A totally non-informative model will yield an AR of zero. The interpretation of AR is that if defaulted firms have been assigned among the highest credit risks before they defaulted, then the model has discriminated properly between the safe and risky firms. The CRI corporate default prediction model achieves high AR scores for all its covered regions and economies, including its good performance.

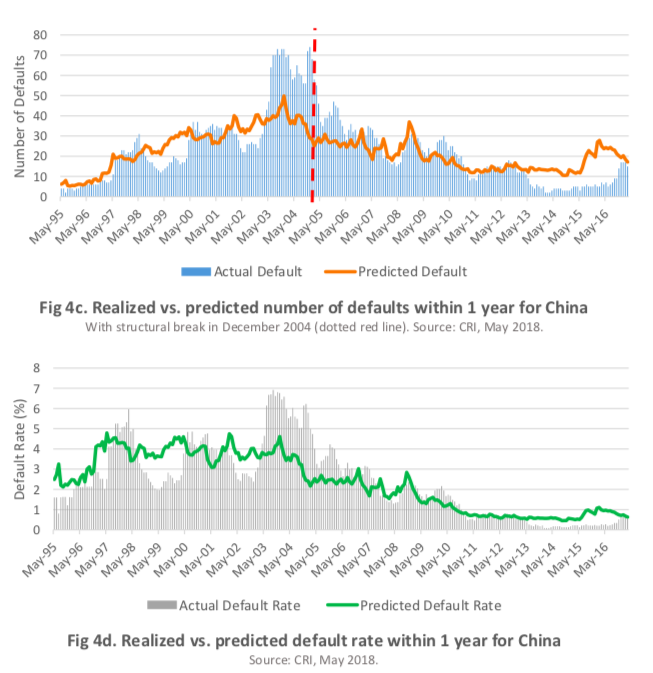
Figure 3 illustrates the AR of the CRI corporate default prediction model for North America, Europe, China for horizons from 1 month to 5 years.



A more straightforward alternative to AR for judging the performance of a default prediction models comparing the actual number of realized defaults against the number of defaults predicted by the model. The following figure 4a to 4d present such comparison using the CRI PD model and realized defaults within one year in the US and China on a monthly frequency.



For the North America calibration group, the CRI PD model introduces a dummy variable on the intercept for financial firms to account for differences that have not been reflected through the other covariates. A structural break in the financial-sector intercept dummy variable has also been introduced to address the dramatic changes after Lehman Brothers’ default in September 2008. Figure 4a through 4d show that the CRI PD model performs well in predicting default numbers and rates for US firms.



For China, the CRI PD model also accounts for a structural break, occurring in December 2004, to reflect the significant drop in default rates after that point. By simply changing the behavior of two parameters (i.e., coefficients for the intercept and DTD level) after December 2004, the performance of the CRI PD model for Chinese firms has significantly improved.

1. Applications

* Predicting Default of a Firm
* Assessing the Aggregate Credit Risk of a Portfolio

Because the CRI computes PD on an individual firm-level basis, the CRI PD of all firms within a specific region and/or sector can easily be aggregated to deliver an overview of the credit environment of that portfolio at a certain point in time.

The aggregate CRI PD well reflects the known financial crisis through rises in PDs in time of crisis. For instance, the 1997 Asian financial crisis particularly affected the credit environment of the ASEAN 5 countries, whereas the internet bubble in 2000-2001 caused wide spread defaults among US firms and financial crisis of 2008-09 severely impacted the US banking sector first before spreading to the rest of the world including the ASEAN 5 countries.

* Other Examples of Application

Aside from assessing the credit risk of a single corporate, and aggregate credit risk of a group or portfolio, CRI PD can be used for a variety of credit management related applications; from benchmarking, to fixed income investment, and credit portfolio management, from model validation to credit research reports.

1. Conclusion

The CRI PD estimates the default risk of publicly listed firms by quantitatively analyzing their financial statements, stock market data and macro-financial factors retrieved from various international data sources. Unlike credit models that utilize letter-grade ratings, the CRI PD is a more granular gauge for credit risk and is available in term structure ranging from 1 month to 5 years. The CRI PD also captures default correlations and can be further aggregated to reflect credit cycles, among others.

The CRI currently provides daily updated PD of over 34,000 active exchange-listed firms globally. The CRI also distributes historical time series of PD of over 33,000 inactive firms due to bankruptcy, corporate consolidation or delisting for other reasons.