NUS-RMI Credit Research Initiative Technical Report

Version: 2018 update 1

Credit Research Initiative[†]
Risk Management Institute
National University of Singapore

This document describes the implementation of the system which the Credit Research Initiative (CRI) at the Risk Management Institute (RMI) of the National University of Singapore (NUS) uses to produce probabilities of default (PD) and actuarial spread (AS). As of this version of the technical report, RMI-CRI covers around 68,000 exchange-listed firms (including delisted ones) in 128 economies around the world (see Table A.1). Of them, over 34,000 firms have sufficient data to release daily updated PD and AS. The PD and AS for all firms are freely available to users who can provide evidence of their professional qualifications to ensure that they will not misuse the data. General users who do not request global access are restricted to a list of 5,000 firms. The individual company PD/AS data, along with aggregate PD/AS at the economy and sector level, can be accessed at `http://rmicri.org`.

The primary goal of this initiative is to drive research and development in the critical area of credit rating systems. As such, a transparent methodology is essential to this initiative. Having the details of the methodology available to everybody means that there is a base from which suggestions and improvements can be made. The objective of this technical report is to provide a full exposition of the CRI system. Readers of this document who have access to the necessary data and who have a sufficient level of technical expertise will be able to implement a similar system on their own. For a full exposition of the conceptual framework of the CRI, see Duan and Van Laere [2012].

The system used by the CRI will evolve as new innovations and enhancements are applied. The main changes to the 2018 technical report and operational implementation of our model are: (1) New smart data launch for the CRI Systematically Important Financial Institution (CriSIFI), (2) New common covariates and some changes in covariates, (3) Changes in parameter estimation, and (4) Expansion of coverage to Qatar.

This technical report reflects such annual updates (1), (2), (3), and (4) until June 12 2018. As of this reference date, the current operational CRI system has been implemented with the model parameters calibrated on June 11 2018 by using available data up to May 31 2018 (henceforth, May calibration). Therefore, all subsequent empirical results (e.g., Tables and Figures in Appendix) are estimated based on May calibration. The latest version of the technical report and addenda are available via the web portal and will include any changes to the system since the publication of this version.

In the remainder of this technical report, the PD model and its computational details will be explained in thorough details. As an application of the model, the computation of AS

[†] For any questions or comments on this article, please contact the CRI Team at rmicri@nus.edu.sg.

and CVI will be discussed in a much concise manner. Wherever no confusion is caused, "the model" refers to the PD model. The sections are organized as follows. Section 1 describes the quantitative model that is currently used to compute the PDs. The model was first described in Duan et al. [2012]. The description includes calibration procedures, which are performed on a monthly basis, and individual firm's PD computations, which are performed on a daily basis.

Section 2 describes the input variables of the model as well as the data used to produce these inputs. This model uses both input variables that are common to all firms in an economy and input variables that are firm-specific. Another critical component in the estimation system is the default data, and this is also described in this section.

While Section 1 provides a broader description of the model, Section 3 describes the implementation details that are necessary for application, given real world issues of, for example, bad or missing data. The specific technical details needed to develop an operational system are also given, including details on the monthly calibration, daily computation of individual firm's PDs, and aggregation of the individual firm's PDs. Distance-to-default (DTD) in a Merton-type model is one of the firm-specific variables. The calculation for DTD is not the standard one, and has been modified to allow a meaningful computation of the DTD for financial firms. While most academic studies on default prediction exclude financial firms from consideration, it is important to include them given that the financial sector is a critical component in every economy. The calculation for DTD is detailed in this section.

Section 4 shows an empirical analysis for those economies that are currently covered. While the analysis shows excellent results in several economies, there is room for improvement in a few others. Basically, all the economies under the CRI coverage adopt the extant variables used in the academic study of US firms in Duan et al. [2012]. As of May 2 2018, we take another step forward by designing new variables to improve default prediction and also start applying variable selection specific to different economies (e.g., China and India). For details, refer to Subsection 2.1. Sections 5 and 6 explain how the CVI and AS are formulated. A detailed theoretical background can be found in Duan [2014]. Section 7 introduces the new CRI product "CriSIFI" aimed at identifying systemic risks of all banks and insurers under the CRI coverage. Section 8 discusses future developments.

1 Model Description

The quantitative model that is currently being used by the CRI is a forward intensity model that was introduced in Duan et al. [2012]. Certain aspects of the model are taken from Duan and Fulop [2013]. This model allows PD forecasts to be made at a range of horizons. In the current CRI implementation of this model, PDs are forecasted from a horizon of one month up to a horizon of five years. At the RMI-CRI website, for every firm, the probability of that firm defaulting within one month, three months, six months, one year, two years, three years, and five years is given. The ability to assess credit quality for different horizons is a useful tool for risk management, credit portfolio management, policy setting, and regulatory purposes, since short- and long-term credit risk profiles can differ greatly depending on a firm's liquidity, debt structures, and other factors.

The forward intensity model is a reduced form model in which the PD is computed as a function of different input variables. These can be firm-specific or common to all firms within an economy. The other category of the default prediction model is the structural model, whereby the corporate structure of a firm is modeled in order to assess the firm's PD.

A similar reduced form model by Duffie et al. [2007] relies on modeling the time series dynamics of the input variables in order to make PD forecasts for different horizons. However, there is little consensus on assumptions for the dynamics of variables such as accounting ratios, and the model output will be highly dependent on these assumptions. In addition, the time series dynamics will be of very high dimension. For example, with the two common variables and two firm-specific variables that Duffie et al. [2007] use a sample of 10,000 firms gives a dimension of the state variables of 20,002.

Given the complexity in modeling the dynamics of variables such as accounting ratios, this model will be difficult to implement if different forecast horizons are required. The key innovation of the forward intensity model is that PD for different horizons can be consistently and efficiently computed based only on the value of the input variables at the time the prediction is made. Thus, the model specification becomes far more tractable.

Fully specifying a reduced form model includes the specification of the function that computes a PD from the input variables. This function is parameterized, and finding appropriate parameter values is called calibrating the model. The forward intensity model can be calibrated by maximizing a pseudo-likelihood function. The calibration is carried out by groups of economies and all firms within a group of economies will use the same parameter values along with each firm's variables in order to compute the firm's PD.

Subsection 1.1 will describe the modeling framework, including the way PDs are computed based on a set of parameter values for the economy and a set of input variables for a firm. Subsection 1.2 explains how the model can be calibrated. Subsection 1.3 details the way parameters are estimated based on the Sequential Monte Carlo (SMC) technique.

1.1 Modeling Framework

While the model can be formulated in a continuous time framework, as done in Duan et al. [2012], an operational implementation requires discretization in time. Since the model is more easily understood in discrete time, the following exposition of the model will begin in a discrete time framework.

Variables for default prediction can have vastly different update frequencies. Financial statement data is updated only once a quarter or even once a year, while market data like stock prices are available at frequencies of seconds. A way of compromising between these two extremes is to have a fundamental time period Δt of one month in the modeling framework. As will be seen later, this does not preclude updating the PD forecasts on a daily basis. This is important since, for example, large daily changes in a firm's stock price can signal changes in credit quality even when there is no change in FS data.

Thus, for the purposes of calibration and subsequently for computing time series of PD, the input variables at the end of each month will be kept for each firm. The input variables associated with the i^{th} firm at the end of the n^{th} month (at time $t = n\Delta t$) is denoted by $X_i(n)$. This is a vector consisting of two parts: $X_i(n) = (W(n), U_i(n))$. Here, $W(n)$ is a vector of variables at the end of month n that is common to all firms in the economy and $U_i(n)$ is a vector of variables specific to firm i.

In the forward intensity model, a firm's default is signaled by a jump in a Poisson process. The probability of a jump in the Poisson process is determined by the intensity of the Poisson process. The forward intensity model draws an explicit dependence of intensities at time periods in the future (that is, forward intensities) to the values of input variables at the time of prediction. With forward intensities, PDs for any forecast horizon can be computed knowing only the values of the input variables at the time of prediction, without needing to simulate future values of the input variables.

There is a direct analogy in interest rate modeling. In spot rate models where dynamics on a short-term spot rate are specified, bond pricing requires expectations on realizations of the short rate. Alternatively, bond prices can be computed directly if the forward rate curve is known.

One issue in default prediction is that firms can exit public exchanges for reasons other than default, such as merge and acquisition (M&A) and OTC. In order to take these other exits into account, defaults and other exits are modeled as two independent Poisson processes, each with their own intensity. While defaults and exits classified as non-defaults are mutually exclusive by definition, the assumption of independent Poisson processes does not pose a problem since the probability of a simultaneous jump in the two Poisson processes is negligible. In the discrete time framework, the probability of simultaneous jumps in the same time interval is non-zero. As a modeling assumption, a simultaneous jump in the same time interval by both the default Poisson process and the non-default type exit Poisson process is considered as a default. In this way, there are three mutually exclusive possibilities during each time interval: survival, default and non-default exit. As with defaults, the forward intensity of the Poisson process for other exits is a function of the input variables. The parameters of this function can also be calibrated.

To further illustrate the discrete framework, the three possibilities for a firm at each time point are diagrammed. Either the firm survives for the next time period Δt, or it defaults within Δt, or it has a non-default exit within Δt. This setup is pictured in Fig. 1. Information about firm i is known up until time $t = m\Delta t$ and the figure illustrates possibilities in the future between $t = (n-1)\Delta t$ and $(n+1)\Delta t$. Here, m and n are integers with $m < n$.

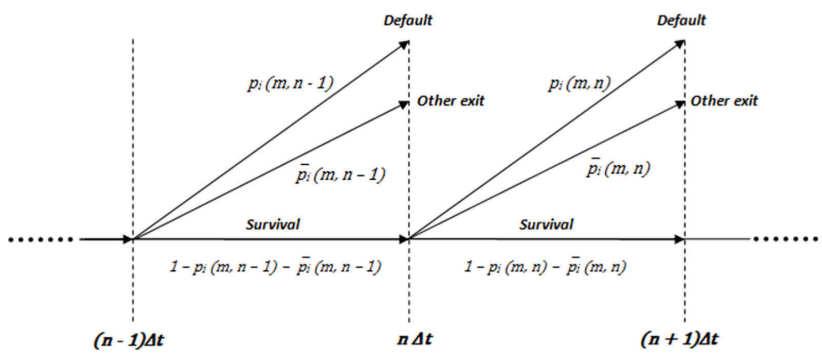

Figure 1: Default-other exit-survival tree for firm i, viewed from time $t = m\Delta t$.

The probabilities of each branch are, for example: $p_i(m, n)$ the conditional probability viewed from $t = m\Delta t$ that firm i will default before $(n+1)\Delta t$, conditioned on firm i surviving up until $n\Delta t$. Likewise, $\bar{p}_i(m, n)$ is the conditional probability viewed from $t = m\Delta t$ that firm i will have a non-default exit before $(n+1)\Delta t$, conditioned on firm i surviving up until $n\Delta t$. It is the modeler's objective to determine $p_i(m, n)$ and $\bar{p}_i(m, n)$, but for now it is assumed that these quantities are known. With the conditional default and other exit probabilities known, the corresponding conditional survival probability of firm i is $1 - p_i(m, n) - \bar{p}_i(m, n)$.

With this diagram in mind, the probability that a particular path will be followed is the product of the conditional probabilities along the path. For example, the probability at time $t = m\Delta t$ of firm i surviving until $(n-1)\Delta t$ and then defaulting between $(n-1)\Delta t$ and $n\Delta t$ is:

$$\mathrm{Prob}_{t=m\Delta t}[\tau_i = n, \tau_i < \bar{\tau}_i] = p_i(m, n-1) \prod_{j=m}^{n-2} \left[1 - p_i(m, j) - \bar{p}_i(m, j)\right]. \tag{1}$$

Here, τ_i is the default time for firm i measured in units of months, $\bar{\tau}_i$ is the other exit time measured in units of months, and the product is equal to 1 if there is no term in the product. The condition $\tau_i < \bar{\tau}_i$ is the requirement that the firm defaults before it has a non-default type of exit. Note that by measuring exits in units of months, if, for example, a default occurs at any time in the interval $[(n-1)\Delta t, n\Delta t]$, then $\tau_i = n$.

Using Eq. (1), cumulative default probabilities can be computed. At $m\Delta t$ the probability of firm i defaulting at or before $n\Delta t$ and not having an other exit before $t = n\Delta t$ is obtained by taking the sum of all of the paths that lead to default at or before $n\Delta t$:

$$\text{Prob}_{t=m\Delta t}[m < \tau_i \leq n, \tau_i < \bar{\tau}_i] = \sum_{k=m}^{n-1} \left\{ p_i(m,k) \prod_{j=m}^{k-1} [1 - p_i(m,j) - \bar{p}_i(m,j)] \right\}. \tag{2}$$

While it is convenient to derive the probabilities given in Eqs. (1) and (2) in terms of the conditional probabilities, expressions for these in terms of the forward intensities need to be found, since the forward intensities will be functions of the input variable $X_i(m)$. The forward intensity for the default of firm i that is observed at time $t = m\Delta t$ for the forward time interval from $t = n\Delta t$ to $(n+1)\Delta t$, is denoted by $h_i(m,n)$, where $m \leq n$. The corresponding forward intensity for a non-default exit is denoted by $\bar{h}_i(m,n)$. Because default is signaled by a jump in a Poisson process, its conditional probability is a simple function of its forward intensity:

$$p_i(m,n) = 1 - \exp[-\Delta t\; h_i(m,n)]. \tag{3}$$

Since joint jumps in the same time interval are assigned as defaults, the conditional other exit probability needs to take this into account:

$$\bar{p}_i(m,n) = \exp[-\Delta t\; h_i(m,n)] \times \{1 - \exp[-\Delta t\; \bar{h}_i(m,n)]\}. \tag{4}$$

The conditional survival probabilities in Eqs. (1) and (2) are computed as the conditional probability that the firm does not default in the period and the firm does not have a non-default exit either:

$$\text{Prob}_{t=m\Delta t}[\tau_i, \bar{\tau}_i > n+1 |\; \tau_i, \bar{\tau}_i > n] = \exp\{-\Delta t[h_i(m,n) + \bar{h}_i(m,n)]\}. \tag{5}$$

It remains to be specified the dependence of the forward intensities on the input variable $X_i(m)$. The forward intensities need to be positive so that the conditional probabilities are non-negative. A standard way to impose this constraint is to specify the forward intensities as exponentials of a linear combination of the input variables:

$$\begin{aligned} h_i(m,n) &= \exp[\beta(n-m) \cdot Y_i(m)], \\ \bar{h}_i(m,n) &= \exp[\bar{\beta}(n-m) \cdot Y_i(m)]. \end{aligned} \tag{6}$$

Here, β and $\bar{\beta}$ are coefficient vectors that are functions of the number of months between the observation date and the beginning of the forward period $(n-m)$, and $Y_i(m)$ is simply the vector $X_i(m)$ augmented by a preceding unit element: $Y_i(m) = (1, X_i(m))$. The unit element allows the linear combination in the argument of the exponentials in Eq. (6) to have a non-zero intercept.

In the current implementation of the forward intensity model in the CRI, the maximum forecast horizon is 60 months (5 years) and there are 16 input variables plus the intercept in general, so there are 60 sets of β and $\bar{\beta}$. While this is a large set of parameters, as will be seen in Subsections 1.2 and 1.3, the calibration is tractable because the default parameters can be calibrated separately from the other exit parameters, and the total number of parameters are greatly reduced after constraining the term-structure of the parameter estimates to be Nelson-Siegel functions.

Before expressing the probabilities in Eqs. (1) and (2) in terms of the forward intensities, a notation H is introduced for the forward intensities so that it becomes clear which parameters the forward intensity depends on:

$$H(\beta(n-m), X_i(m)) = \exp[\beta(n-m) \cdot Y_i(m)]. \tag{7}$$

This is the forward default intensity. The corresponding notation for other exit forward intensities is then just $H(\bar{\beta}(n-m), X_i(m))$. So, the probability in Eq. (1) is expressed in terms of the forward intensities, using Eq. (3) as the conditional default probability and Eq. (5) as the conditional survival probability:

$$\begin{aligned}
&\text{Prob}_{\,t=m\Delta t}[\tau_i = n, \tau_i < \bar{\tau}_i] \\
&\quad = \{1-\exp[-\Delta t\, H(\beta(n-1-m), X_i(m))]\} \\
&\qquad \times \prod_{j=m}^{n-2} \exp\left\{-\Delta t\, [H(\beta(j-m), X_i(m)) + H(\bar{\beta}(j-m), X_i(m))]\right\} \\
&\quad = \{1-\exp[-\Delta t\, H(\beta(n-m-1), X_i(m))]\} \\
&\qquad \times \exp\left\{-\Delta t \sum_{j=m}^{n-2} \left[H(\beta(j-m), X_i(m)) + H(\bar{\beta}(j-m), X_i(m))\right]\right\}.
\end{aligned} \tag{8}$$

This probability will be relevant in the next part during the calibration. The cumulative default probability given in Eq. (2) in terms of the forward intensities is then:

$$\begin{aligned}
&\text{Prob}_{\,t=m\Delta t}[m < \tau_i \le n, \tau_i < \bar{\tau}_i] \\
&\quad = \sum_{k=m}^{n-1} \Bigg\{ \{1-\exp[-\Delta t\, H(\beta(k-m), X_i(m))]\} \\
&\qquad \times \exp\left\{-\Delta t \sum_{j=m}^{k-1} [H(\beta(j-m), X_i(m)) + H(\bar{\beta}(j-m), X_i(m))]\right\}\Bigg\}.
\end{aligned} \tag{9}$$

This formula is used to compute the main output of the CRI: an individual firm's PD within various time horizons. The β and $\bar{\beta}$ parameters are obtained when the firm's economy is calibrated, and using those together with the firm's input variables yields the firm's PD.

1.2 Pseudo-Likelihood Function

The empirical data set used for calibration can be described as follows. For the economy as a whole, there are N end of month observations, indexed as $n = 1, \ldots, N$. Of course, not all firms will have observations for each of the N months as they may start later than the start of the economy's data set or they may exit before the end of the economy's data set. There are a total of I firms in the economy, and they are indexed as $i = 1, \ldots, I$. As before, the input variables for the i^{th} firm in the n^{th} month is $X_i(n)$. The set of all observations for all firms is denoted by X.

In addition, the default times τ_i and non-default exit times $\bar{\tau}_i$ for the i^{th} firm are known if the default or other exit occurs after time $t = \Delta t$ and at or before $t = N\Delta t$. The possible values for τ_i and $\bar{\tau}_i$ are integers between 2 and N, inclusive. If a firm exits before the month end, then the exit time is recorded as the first month end after the exit. If the firm does not exit before $t = N\Delta t$, then the convention can be used that both of these values are infinite. If the firm has a default type of exit within the data set, then $\bar{\tau}_i$ can be considered as infinite. If instead the firm has a non-default type of exit within the data set, then τ_i can be considered as infinite. The set of all default times and non-default exit times for all firms is denoted by τ and $\bar{\tau}$, respectively. The first month in which firm i has an observation is denoted by t_{0i}. Except for cases of missing data, these observations continue until the end of the data set if the firm never exits. If the firm does exit, the last needed input variable $X_i(n)$ is for $n = \min(\tau_i, \bar{\tau}_i) - 1$.

The calibration of the β and $\bar{\beta}$ parameters is done by maximizing a pseudo-likelihood function. The function to be maximized violates the standard assumptions of likelihood functions, but Appendix A in Duan et al. [2012] derives the large sample properties of the pseudo-likelihood function.

In formulating the pseudo-likelihood function, the assumption is made that the firms are conditionally independent of each other. In other words, correlations arise naturally from shared common factors $W(n)$ and any correlations between different firms' firm-specific variables. With this assumption, the pseudo-likelihood function for the horizon of ℓ months, a set of parameters β and $\bar{\beta}$ and the data set $(\tau, \bar{\tau}, X)$ is:

$$\mathcal{L}_\ell(\beta, \bar{\beta}; \tau, \bar{\tau}, X) = \prod_{m=1}^{N-1} \prod_{i=1}^{I} P_{\min(N-m,\ell)}(\beta, \bar{\beta}; \tau_i, \bar{\tau}_i, X_i(m)). \tag{10}$$

Here, $P_{\min(N-m,\ell)}(\beta, \bar{\beta}; \tau_i, \bar{\tau}_i, X_i(m))$ is a probability for the i^{th} firm, with the nature of the probability depending on what happens to the firm during the period from month m to month $m + \min(N - m, \ell)$. This is defined as:

$$\begin{aligned}
&P_\ell(\beta, \bar{\beta}; \tau_i, \bar{\tau}_i, X_i(m)) \\
&\quad = 1_{\{t_{0i} \le m, \min(\tau_i, \bar{\tau}_i) > m+\ell\}} \\
&\qquad \times \exp\left\{ -\Delta t \sum_{j=0}^{\ell-1} [H(\beta(j), X_i(m)) + H(\bar{\beta}(j), X_i(m))] \right\} \\
&\qquad + 1_{\{t_{0i} \le m, \tau_i \le \bar{\tau}_i, \tau_i \le m+\ell\}} \times \{1 - \exp[-\Delta t\, H(\beta(\tau_i - m - 1), X_i(m))]\} \\
&\qquad \times \exp\left\{ -\Delta t \sum_{j=0}^{\tau_i - m - 2} [H(\beta(j), X_i(m)) + H(\bar{\beta}(j), X_i(m))] \right\} \\
&\qquad + 1_{\{t_{0i} \le m, \bar{\tau}_i \le \tau_i, \bar{\tau}_i \le m+\ell\}} \times \{1 - \exp[-\Delta t\, H(\bar{\beta}(\bar{\tau}_i - m - 1), X_i(m))]\} \\
&\qquad \times \exp[-\Delta t H(\beta(\tau_i - m - 1), X_i(m))] \\
&\qquad \times \exp\left\{ -\Delta t \sum_{j=0}^{\bar{\tau}_i - m - 2} [H(\beta(j), X_i(m)) + H(\bar{\beta}(j), X_i(m))] \right\} \\
&\qquad + 1_{\{t_{0i} > m\}} + 1_{\{\min(\tau_i, \bar{\tau}_i) \le m\}}.
\end{aligned} \tag{11}$$

In other words, if the i^{th} firm survives from the observation time at month m for the full horizon ℓ until at least $m + \ell$, then the probability is the model-based survival probability for this period. This is the first term in Eq. (11). The second term handles the cases where the firm has a default within the horizon, in which case the probability is the model-based probability of the firm defaulting at the month that it ends up defaulting, as given in Eq. (8). The third term handles the cases where the firm has a non-default exit within the horizon, in which case the probability is the model-based probability of the firm having a non-default type exit at the month that the exit actually does occur. The expression for this probability uses the conditional non-default type exit probability given in Eq. (4). The final two terms handle the cases where the firm is not in the data set at month m - either the first observation for the firm is after m or the firm has already exited. A constant value is assigned in this case so that this firm will not affect the maximization at this time point.

The pseudo-likelihood function given in Eq. (10) can be numerically maximized to give estimates for the coefficients β and $\bar{\beta}$. Notice though that the sample observations for the pseudo-likelihood function are overlapping if the horizon is longer than one month. For example, when $\ell = 2$, default over the next two periods from month m is correlated to default over the next two periods from month $m + 1$ due to the common month in the two sample observations. However, in Appendix A of Duan et al. [2012], the maximum pseudo-likelihood estimator is shown to be consistent, in the sense that the estimators converge to the "true" parameter value in the large sample limit.

Notice though that each of the terms in Eq. (11) can be written as a product of terms containing only β and terms containing only $\bar{\beta}$. This will allow separate maximizations with respect to β and with respect to $\bar{\beta}$, that is, the defaults and other exits.

The β and $\bar{\beta}$ specific versions of Eq. (11) are:

$$
\begin{aligned}
&P_\ell^{\beta}(\beta;\tau_i,\bar{\tau}_i,X_i(m)) \\
&\quad = 1_{\{t_{0i}\le m,\min(\tau_i,\bar{\tau}_i)>m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\ell-1} H(\beta(j),X_i(m))\right\} \\
&\qquad + 1_{\{t_{0i}\le m,\tau_i\le\bar{\tau}_i,\tau_i\le m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\tau_i-m-2} H(\beta(j),X_i(m))\right\} \\
&\qquad\quad \times\{1-\exp[-\Delta t\; H(\beta(\tau_i-m-1),X_i(m))]\} \\
&\qquad + 1_{\{t_{0i}\le m,\bar{\tau}_i\le\tau_i,\bar{\tau}_i\le m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\bar{\tau}_i-m-2} H(\beta(j),X_i(m))\right\} \\
&\qquad\quad \times\exp[-\Delta t\; H(\beta(\tau_i-m-1),X_i(m))] \\
&\qquad + 1_{\{t_{0i}>m\}} + 1_{\{\min(\tau_i,\bar{\tau}_i)\le m\}},
\end{aligned}
$$

$$
\begin{aligned}
&P_\ell^{\bar{\beta}}(\bar{\beta};\tau_i,\bar{\tau}_i,X_i(m)) \\
&\quad = 1_{\{t_{0i}\le m,\min(\tau_i,\bar{\tau}_i)>m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\ell-1} H(\bar{\beta}(j),X_i(m))\right\} \\
&\qquad + 1_{\{t_{0i}\le m,\tau_i\le\bar{\tau}_i,\tau_i\le m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\tau_i-m-2} H(\bar{\beta}(j),X_i(m))\right\} \\
&\qquad + 1_{\{t_{0i}\le m,\bar{\tau}_i\le\tau_i,\bar{\tau}_i\le m+\ell\}} \exp\left\{-\Delta t\sum_{j=0}^{\bar{\tau}_i-m-2} H(\bar{\beta}(j),X_i(m))\right\} \\
&\qquad\quad \times\{1-\exp[-\Delta t\; H(\bar{\beta}(\bar{\tau}_i-m-1),X_i(m))]\} \\
&\quad + 1_{\{t_{0i}>m\}} + 1_{\{\min(\tau_i,\bar{\tau}_i)\le m\}}.
\end{aligned}
\tag{12}
$$

Then, the β and $\bar{\beta}$ specific versions of the pseudo-likelihood function are given by:

$$
\begin{aligned}
\mathcal{L}_\ell^{\beta}(\beta;\tau,\bar{\tau},X) &= \prod_{m=1}^{N-\ell}\prod_{i=1}^{I} P_\ell^{\beta}(\beta;\tau_i,\bar{\tau}_i,X_i(m)) \\
\mathcal{L}_\ell^{\bar{\beta}}(\bar{\beta};\tau,\bar{\tau},X) &= \prod_{m=1}^{N-\ell}\prod_{i=1}^{I} P_\ell^{\bar{\beta}}(\bar{\beta};\tau_i,\bar{\tau}_i,X_i(m)).
\end{aligned}
\tag{13}
$$

With the definitions given in Eqs. (12) and (13), it can be seen that:

$$
\mathcal{L}_\ell(\beta,\bar{\beta};\tau,\bar{\tau},X) = \mathcal{L}_\ell^{\beta}(\beta;\tau,\bar{\tau},X)\mathcal{L}_\ell^{\bar{\beta}}(\bar{\beta};\tau,\bar{\tau},X). \tag{14}
$$

Thus, $\mathcal{L}_\ell^{\beta}$ and $\mathcal{L}_\ell^{\bar{\beta}}$ can be separately maximized to find their respective parameters. Subsection 1.3 will further explain how the optimum parameters can be estimated.

1.3 Parameter Estimation

Previously, the CRI system produced default predictions to a horizon of two years (CRI [2012]). An extension of the forecast horizon has been implemented as of the PD released on April 1 2013. With this update, horizons of up to five years are now being computed. Technically speaking, horizons of arbitrary length can be calculated.

This extension to a five-year horizon is done by constraining the term-structure of the parameter estimates to be Nelson-Siegel (Nelson and Siegel [1987]; hereafter NS) functions of the forward-starting time. Horizon-specific parameters β and $\bar{\beta}$ can be obtained from the continuous NS function by using the forward prediction horizon as an input. The term-structures are further constrained so that the effect of risk factors on the forward intensity goes to zero as the horizon increases. This allows tractable and parsimonious extrapolations for horizons beyond five years.

The parameter estimation for the NS functions is based on a new numerical method (a pseudo-Bayesian SMC technique) developed by Duan and Fulop [2013]. The remainder of this section details the new parameter estimation. Subsection 1.3.1 describes the parameterization of the parameters by NS functions. Subsection 1.3.2 explains how a structural break applies to the CRI-PD model parameters for the North America calibration group and Chinese firms. Subsection 1.3.3 gives an overview of the SMC method that is used to estimate the NS functions. Subsection 1.3.4 details the calculation of the confidence intervals for the parameter estimation, and Subsection 1.3.5 describes how the parameters can be re-estimated given new data or updates of old data.

1.3.1 Smoothed parameters

Duan et al. [2012] formulate the forward intensity model in which the forward default intensity for a firm is a function of a number of covariates. The forward default intensities for different forward starting periods are computed using different sets of parameters.

In Duan et al. [2012], the sets of parameters are estimated separately for each forward starting time. Parameters at different forward starting times that are associated with each covariate can be approximated by a function of the forward starting time using NS type term structure functions. Duan et al. [2012] show that this approximation by NS functions does not negatively affect prediction performance. The CRI implementation follows Duan and Fulop [2013] to impose the functional restriction during the estimation as opposed to the method used in Duan et al. [2012] of fitting the curve after parameter estimates have been obtained. This is done for two reasons.

First, it will significantly reduce the number of parameters. For example, using 16 covariates for forward default intensities up to 60 months would require a joint estimation of $17 \times 60 = 1020$ parameters. Here, 17 comes from adding an intercept to the intensity function with 16 covariates. If the coefficients corresponding to each covariate are represented by the NS function of 4 parameters, there will be at most $17 \times 4 = 68$ parameters. In fact, there will be fewer parameters as some of the NS parameters will be constrained to zero.

Second, the NS function will allow extrapolation. For example, the 17 NS functions estimated with predictions up to 60 months can be used for prediction, say, over 72 months.

The NS function with four free parameters is:

$$r(t; \varrho_0, \varrho_1, \varrho_2, d) = \varrho_0 + \varrho_1 \frac{1 - \exp(-t/d)}{t/d} + \varrho_2 \left[\frac{1 - \exp(-t/d)}{t/d} - \exp(-t/d) \right], \tag{15}$$

where t is the forecast horizon (measured in years). In the CRI implementation, the horizon is 60 months (5 years) so that t ranges from 0 to 59/12. Once the four NS parameters are estimated, individual horizon-specific parameters β and $\bar{\beta}$ are obtained from the NS function r using the forecast horizon as input to the NS function. In our current implementation with forecast horizons extending to 60 months (5 years), 120 sets of month specific β and $\bar{\beta}$ are obtained. For all covariates, the restriction $d > 0$ is imposed so that the functions converge to a value for large t. This formulation will be used for forward intensities for both defaults and other types of exit.

For the coefficients of all stochastic covariates, the long-run level ϱ_0 is restricted to zero, because the current value of a stochastic covariate should be uninformative of default or other exits when the forward starting time goes to infinity. In other words, the coefficient of such a stochastic covariate should approach zero when t goes to infinity.

The intercept of the forward intensity function is of course non-stochastic. Thus, ϱ_0 can have non-zero values for the intercept. With these restrictions on the NS parameters, take the example of 16 covariates and an intercept, there will be a total of $16 \times 3 + 1 \times 4 = 52$ parameters, provided that the calibration group does not carry a structural break.

In the CRI implementation, the NS function is further constrained to be non-positive for certain covariates: liquidity level and trend, and profitability level and trend. Refer to Section 2 for descriptions of these covariates.

For China, we have 15 input variables (an intercept plus 14 covariates) due to the different variable selection specific to the economy (see Subsection 2.1). In addition, we further revise the parameter estimation for the North America calibration group and Chinese firms. For details, refer to Subsection 1.3.2.

1.3.2 Structural break

The North America calibration group (the US and Canada) has incorporated the following two specific changes. First, we include a dummy variable on the intercept for financial firms to account for differences that have not been duly reflected through other covariates. Second, we apply a structural break to this financial-sector intercept dummy to address the change in September 2008 after Lehman Brothers defaulted.

The structural break for the North America calibration group is treated as an impulse response. The key is to allow the different rates of transition, characterized by $\widetilde{\alpha}_1(\tau) > 0$ and $\widetilde{\alpha}_2(\tau) > 0$, before and after the break point t_0 (September 2008), respectively. Before t_0, for example, the coefficient for the financial–sector intercept dummy, $\beta(t, \tau; t_0)$, has the form:

$$\beta(t, \tau; t_0) = \widetilde{\beta}(\tau) + \widetilde{\gamma}(\tau) \times \frac{1}{1 + e^{-\widetilde{\alpha}_1(\tau)(t - t_0)}},$$

where t denotes the default prediction time, and τ denotes a forward starting time ranging from 0 (1 month) to 59/12 (5 years). $\widetilde{\alpha}_1(\tau)$, $\widetilde{\beta}(\tau)$, and $\widetilde{\gamma}(\tau)$ are characterized by the NS function in Eq. (15). After t_0, the coefficient for the financial–sector intercept dummy is governed by $\widetilde{\alpha}_2(\tau)$ instead of $\widetilde{\alpha}_1(\tau)$:

$$\beta(t, \tau; t_0) = \widetilde{\beta}(\tau) + \widetilde{\gamma}(\tau) \times \frac{1}{1 + e^{-\widetilde{\alpha}_2(\tau)(t_0 - t)}}.$$

Therefore, $\beta(t, \tau; t_0)$ moves from $\widetilde{\beta}(\tau)$ to $\widetilde{\beta}(\tau) + 1/2\widetilde{\gamma}(\tau)$ as t advances toward t_0, and reverts back to $\widetilde{\beta}(\tau)$ as t goes past t_0.

Our treatment on Chinese firms differs from that for the North American calibration group in two aspects. First, we apply a structural break to both the intercept and the DTD level. Second, we model the structural break by a step function allowing for different rates of transition to and away from the break point. As implemented earlier, the treatment is the same for intercept term and the coefficient for the DTD level, but the transition rates are different. Here, we describe generically for one of these two structural breaks. Before t_0 (December 2004), $\beta(t, \tau; t_0)$ has the following form:

$$\beta(t, \tau; t_0) = \widetilde{\beta}(\tau) + \widetilde{\gamma}(\tau) \times \frac{1}{1 + e^{-\widetilde{\alpha}_1(\tau)(t - t_0)}},$$

After t_0, the two variables are governed by $\widetilde{\alpha}_2(\tau)$:

$$\beta(t, \tau; t_0) = \widetilde{\beta}(\tau) + \widetilde{\gamma}(\tau) \times \frac{1}{1 + e^{-\widetilde{\alpha}_2(\tau)(t - t_0)}}.$$

Therefore, $\beta(t,\tau;t_0)$ smoothly transits from $\widetilde{\beta}(\tau)$ to $\widetilde{\beta}(\tau)+1/2\widetilde{\gamma}(\tau)$ as t moves toward t_0, and then continues to $\widetilde{\beta}(\tau)+\widetilde{\gamma}(\tau)$ as t moves beyond t_0.

1.3.3 Parameter estimation by SMC

Reliably estimating a system involving 52 parameters for 16 covariates and an intercept presents a numerical challenge. Moreover, the number of parameters can be greater than 52 if there are more than 16 covariates or structural breaks. The CRI implementation follows Duan and Fulop [2013] who use the SMC pseudo-Bayesian method for estimation and self-normalized statistics for inference.

Due to decomposability, the analysis can be performed separately on the forward default and other exit intensities. The data in the CRI implementation are refreshed with monthly frequency, and the sample likelihood used in estimation relies on default predictions running from 1 month to 60 months with a one month increment. Naturally, default prediction is subject to data availability. Towards the end of the period with available data, the prediction horizon naturally decreases and stops at one-month predictions.

The following exposition closely follows the appendix in Duan and Fulop [2013]. It is important to note that the CRI implementation uses the model described in Duan and Fulop [2013], which does not contain any latent frailty or partial conditioning variable, and hence is technically much simpler in parameter estimation. For example, there is no nonlinear filtering problem.

According to the current modeling framework, where for a particular economy there are N end of month observations, the input variables of the ith firm in the mth month is given by $X_i(m)$. Let θ denote a set of NS parameters and ℓ denote the forecast horizon ($\ell=60$). Then the pseudo-likelihood function at step m, denoted by $\mathcal{L}_{m,\min(N-m,\ell)}(\theta)$, takes the form:

$$\mathcal{L}_{m,\min(N-m,\ell)}(\theta)=\prod_{i=1}^{I}P_{\min(N-m,\ell)}\left(\beta(\theta),\bar{\beta}(\theta);\tau_i,\bar{\tau}_i,X_i(m)\right), \tag{16}$$

where I is the number of firms, $\beta(\theta)$ and $\bar{\beta}(\theta)$ are the default and other exit coefficient vectors from Eq. (6) generated from the NS functions with parameter θ, respectively. One may notice that $\mathcal{L}_{m,\min(N-m,\ell)}(\theta)$ is one of the terms in the outer-most product in Eq. (10).

Let $\pi(\theta)$ denote the prior. Following the notation from Section 1.1, consider the following pseudo-posterior distribution at time n after one makes the ℓ-period prediction:

$$\gamma_n(\theta)\propto\prod_{m=1}^{n-1}\mathcal{L}_{m,\min(N-m,\ell)}(\theta)\pi(\theta),\ \text{for } n=2,\dots,N, \tag{17}$$

In the CRI implementation, $\pi(\theta)$ is set to 1 (i.e., a uniform or improper prior) instead of the previous normal/truncated normal priors. This revision frees the estimation algorithm from needing an ad hoc prior belief to start the process. Despite this change, the estimation results remain qualitatively similar, reflecting the fact that our dataset is quite large and the prior's effect is only marginal.

One can apply the sequential batch-resampling routine of Chopin [2002] together with tempering steps as in Del Moral et al. [2006] to advance the system. For each n, this procedure yields a weighted sample of K particles, $(\theta^{(k,n)},w^{(k,n)})$ for $k=1,\dots,K$, whose empirical distribution function will converge to $\gamma_n(\theta)$ as K increases. In the following paragraphs, the superscript k denotes the particle index. Note that in the CRI implementation, K=1,000.

Initialization: To provide the initial particle cloud from which the algorithm can start, an initial random sample from the normal distribution is drawn ($\theta^{(k,0)}\sim\mathcal{N}(\boldsymbol{\mu},\Sigma)$, $w^{(k,0)}=1/K$). Of course, the support of the normal distribution must contain the true parameter value θ_0. In the CRI implementation, $\boldsymbol{\mu}$ and σ are chosen based on cumulative knowledge on parameters' locations and dispersions to speed up optimization.

***Recursions and defining the tempering sequence*:** Assume there is a particle cloud $\left(\theta^{(k,n)}, w^{(k,n)}\right)$ whose empirical distribution represents $\gamma_n(\theta)$. Then, a cloud representing $\gamma_{n+1}(\theta)$ will be reached by combining importance sampling and the Markov Chain Monte Carlo (MCMC) steps. Sometimes moving directly from $\gamma_n(\theta)$ to $\gamma_{n+1}(\theta)$ is too ambitious as the two distributions are too far from each other. This will be reflected in highly variable importance weights if one resorts to direct importance sampling. Hence, following Duan and Fulop [2013] which in turn followed Del Moral et al. [2006], a tempered bridge is built between the two densities and the particles are evolved through the resulting sequence of densities. In particular, assume that at time $n+1$, there are P_{n+1} intermediate densities:

$$\overline{\gamma}_{n+1,p}(\theta) \propto \gamma_n(\theta)\mathcal{L}^{\xi_p}_{n,\min(N-n,\ell)}(\theta), \text{ for } p = 0, \ldots, P_{n+1}. \tag{18}$$

This construction defines an appropriate bridge: $\xi_0 = 0$ so that $\overline{\gamma}_{n+1,0}(\theta) = \gamma_n(\theta)$, and $\xi_{P_{n+1}} = 1$ so that $\overline{\gamma}_{n+1,P_{n+1}}(\theta) = \gamma_{n+1}(\theta)$. For p between 0 and P_{n+1}, ξ_p is chosen from a grid of points to evenly distribute the weights, as described below. A particle cloud representing $\overline{\gamma}_{n+1,0}(\theta)$ can be initialized as $(\overline{\theta}^{(k,n+1,0)}, \overline{w}^{(k,n+1,0)}) = (\theta^{(k,n)}, w^{(k,n)})$. Then, for $p = 1, \ldots, P_{n+1}$ the sequence proceeds as follows:

- *Reweighting Step:* At the beginning of each tempering step, p, a reweighting procedure is run:

$$\overline{w}^{(k,n+1,p-1)} \times \mathcal{L}^{\xi_p - \xi_{p-1}}_{n,\min(N-n,\ell)}(\overline{\theta}^{(k,n+1,p)}), \tag{19}$$

  where ξ_p is chosen to ensure that a minimum effective sample size (ESS) is maintained, where ESS is defined as

$$\text{ESS} = \frac{\left(\sum_{k=1}^{K} \overline{w}^{(k,n+1,p)}\right)^2}{\sum_{k=1}^{K}\left(\overline{w}^{(k,n+1,p)}\right)^2}. \tag{20}$$

  The newly adopted minimum ESS is 25% of the sample size, which equals 250 with the CRI's use of the SMC sample for 1,000 parameter particles. This is done by a grid search, where the ESS is evaluated at a grid of candidate values for ξ_p. The one that produces the ESS that is larger than and closest to 250 is chosen. By changing the criterion from 500 to 250, bigger steps for ξ_p are taken to speed the algorithm without adversely affecting the quality of the estimation result.

  In order to arrive at a representation of $\overline{\gamma}_{n+1,p}(\theta)$, the particles representing $\overline{\gamma}_{n+1,p-1}(\theta)$ and the importance sampling principle can be used. This leads to:

$$\overline{\theta}^{(k,n+1,p)} = \overline{\theta}^{(k,n+1,p-1)}, \tag{21}$$

$$\begin{aligned}\overline{w}^{(k,n+1,p)} &= \overline{w}^{(k,n+1,p-1)} \times \frac{\overline{\gamma}_{n+1,p}(\overline{\theta}^{(k,n+1,p)})}{\overline{\gamma}_{n+1,p-1}(\overline{\theta}^{(k,n+1,p)})} \\ &= \overline{w}^{(k,n+1,p-1)} \times \mathcal{L}^{\xi_p - \xi_{p-1}}_{n,\min(N-n,\ell)}(\overline{\theta}^{(k,n+1,p)}).\end{aligned} \tag{22}$$

  To avoid particle impoverishment in sequential importance sampling where most of the weights are concentrated in a small number of particles, a resample-move step is run.

- *Resampling Step:* The particles are resampled proportional to their weights. If $I^{(k,n+1,p)} \in (1, \ldots, K)$ are particle indices sampled proportional to $\overline{w}^{(k,n+1,p)}$, the equally weighted particles are obtained as

$$\overline{\theta}^{(k,n+1,p)} = \overline{\theta}^{(I^{(k,n+1,p)},n+1,p)}, \tag{23}$$

$$\overline{w}^{(k,n+1,p)} = \frac{1}{K}. \tag{24}$$

- *Move Step:* Each particle is passed through a Markov kernel $\mathcal{K}_{n+1,p}(\overline{\theta}^{(k,n+1,p)}, \cdot)$ that leaves $\overline{\gamma}_{n+1,p}(\theta)$ invariant, typically a Metropolis-Hastings kernel:

  1. Propose $\theta^{*(k)} \sim \mathcal{Q}_{n+1,p}\left(\cdot \;\middle|\; \overline{\theta}^{(k,n+1,p)} \right)$.
  2. Compute the acceptance rate α, where:

  $$\alpha = \min\left(1, \frac{\overline{\gamma}_{n+1,p}(\theta^{*(k)})\mathcal{Q}_{n+1,p}(\overline{\theta}^{(k,n+1,p)} \mid \theta^{*(k)})}{\overline{\gamma}_{n+1,p}(\overline{\theta}^{(k,n+1,p)})\mathcal{Q}_{n+1,p}(\theta^{*(k)} \mid \overline{\theta}^{(k,n+1,p)})}\right). \tag{25}$$

  3. With probability α, set $\overline{\theta}^{(k,n+1,p)} = \theta^{*(k)}$, otherwise keep the old particle.

This step will enrich the support of the particle cloud while conserving its distribution. If the particle set is a poor representation of the target distribution, the move step can also help adjust the location of the support. Crucially, given the importance of the sampling setup, the proposal distribution $Q_{n+1,p}(\cdot \mid \overline{\theta}^{(k,n+1,p)})$ can be adapted using the existing particle cloud.

In the CRI implementation, we define three (or four) NS parameters corresponding to each covariate as one block. A mixture distribution is designed to combine with equal probabilities: (1) a block independent normal distribution using the means and the standard deviations derived from the existing particle set, and (2) a random walk proposal based on a scaled-down covariance matrix used in the block independent proposal; that is,

$$\theta^{*(k)} \sim \frac{1}{2}\mathcal{N}(\boldsymbol{\mu}, \Sigma) + \frac{1}{2}\mathcal{N}\left(\overline{\theta}^{(k,n+1,p)}, \Sigma^*\right),$$

where $\boldsymbol{\mu}$ is the sample mean vector of $\overline{\theta}^{(k,n+1,p)}$ and Σ is the covariance matrix with a block diagonal structure, i.e., the covariances across blocks are all zero. $\sigma^{*2}_{i,j}$, which is the (i, j)–th element of Σ^*, is set to be $(0.2\sigma_{i,j})^2$ (the (i, j)–th element of Σ), to propose around the original values. Mixing the independent and random walk proposals can effectively boost the support (i.e., a higher ESS) by offering local alternatives to those parameters with already high likelihood, especially when there exists discrepancies between the true distribution and its approximating normal distribution.

Moreover, we do not propose to replace an entire parameter particle, and implement a random block proposal. For each particle, say, comprising sixteen blocks (i.e., covariates), we randomly select a random number of blocks (from five to ten) and only propose new values for the selected blocks, while keeping the remaining blocks at their original values. This design can increase the acceptance rate and still offer rich enough replacements. To ensure a good replacement for every block, we perform multiple such Metropolis-Hastings steps each time until the accumulated acceptance rate exceeds 100% and the ESS reaches at least 75% of sample size.

Finally, proposed particles must satisfy some pre-defined constraints. First, the NS parameter d must be positive. Second, particles must produce an increasing or decreasing structure of the NS function for the first five months in order to ensure the smoothness of the term structure of the forward intensity parameters. Third, the coefficients for some covariates, such as the level and trend of liquidity, are required to be non–positive over all forward starting times.

Using the mixture proposal creates a minor complication. The sampler for the truncated values does not carry the same norming constant due to the inclusion of the random walk proposal so that it cannot be ignored in the importance weight. To address the issue, we treat those sampled parameters violating the above mentioned constraints as if there were legitimate particles, but assign the likelihood $\overline{\gamma}_{n+1,p}(\theta^{*(k)})$ of any such proposed particle a value of 0. In short, such particles will never be accepted.

***Final tempering step*:** When $p = P_{n+1}$ is reached (i.e., ξ_p reaches 1), a representation of $\gamma_{n+1}(\theta)$ is:

$$(\theta^{(k,n+1)}, w^{(k,n+1)}) = (\overline{\theta}^{(k,n+1,P_{n+1})}, \overline{w}^{(k,n+1,P_{n+1})}). \tag{26}$$

Additional Metropolis–Hastings moves are performed until the accumulated acceptance rate exceeds 200% instead of 100% at the prior steps. This is to improve the final quality of the SMC sample of parameter particles in representing the target distribution.

***Re–initialization*:** Recall that our SMC approach is the expanding–data SMC technique according to the classification in Duan and Fulop [2013]. Although the expanding data approach is more computationally efficient, we noticed that approximation errors may sometimes get accumulated after repeatedly updating the SMC parameter particle set by adding data one month at a time. We thus introduce a parameter re-initialization every 10 sequential updating time steps to remove the potentially accumulated approximation errors. Re-initialization is the same as the initialization at the beginning of the SMC, except that the relevant means and variances-covariances are computed with the updated SMC parameter particle set so that re-initialization can take advantage of updated information on the sampling distribution.

1.3.4 Statistical inference

The full sample size has N time series data points, but one can only make default prediction at $N-1$ time points; for example, at time point 2, the data is only available for making one-period default prediction at time point 1. Denote the pseudo-posterior mean of the parameter of the whole sample by $\hat{\theta}_N$. And for $n = 2, \cdots, N$,

$$\hat{\theta}_n = \frac{1}{\sum_{k=1}^{K} w^{(k,n)}} \sum_{k=1}^{K} w^{(k,n)} \theta^{(k,n)}. \tag{27}$$

Note that $(\overline{\theta}^{(k,n+1,0)}, \overline{\omega}^{(k,n+1,0)}) = (\theta^{(k,n)}, \omega^{(k,n)})$ is not a true posterior because the likelihood function in Eq. (17) is not a true likelihood function. Thus, it cannot directly provide valid Bayesian inference. But following Duan and Fulop [2013] - which is in turn based on Shao's self-normalized statistic (Shao [2010]) - inference can be performed using the t-like statistic in the full-sample run. To test, for example, the hypothesis of the kth element of $\overline{\theta}^{(k,n+1,p)} = \overline{\theta}^{(I^{(k,n+1,p)},n+1,p)}$, denoted by $\overline{\omega}^{(k,n+1,p)} = \frac{1}{K}$, equal to a, one has:

$$t^* = \frac{\sqrt{N-1}\left(\hat{\theta}_N^{(k)} - a\right)}{\sqrt{\hat{\delta}_{k,N}}} \xrightarrow{d} \frac{W(1)}{\left[\int_0^1 (W(r) - rW(1))^2 dr\right]^{1/2}}, \tag{28}$$

where $W(r)$ is a Wiener process, $\hat{\delta}_{k,N}$ is the kth diagonal element of $\hat{C}_N$, and

$$\hat{C}_N = \frac{1}{(N-1)^2} \sum_{n=2}^{N} n^2 (\hat{\theta}_n - \hat{\theta}_N)(\hat{\theta}_n - \hat{\theta}_N)'. \tag{29}$$

The statistical inference on the structural break parameters are again based on Shao's self-normalized statistic (see Subsection 1.3.2). Since the parameters in connection with the structural break cannot be identified using the data before the break point, the sequence of parameter estimates used in Shao's self-normalized statistic can only start from the break point onward. In the CRI implementation, all parameter estimates, break or non-break related, start from the break point. Denote by T the endpoint of the data set and t_0 again the structural break point. The number of points in the sequence, N, used to compute the norming matrix and the confidence intervals (see Eq. (29)) therefore equals $T - t_0 + 1$.

The right-hand-side random variable for t^* in Eq. (28) does not have a known distribution, but can be easily simulated. Kiefer et al. [2000] reported that the 95% quantile is 5.374 and the 97.5% quantile is 6.811. These values can also be used to set up confidence intervals.

1.3.5 Periodic updating

In reality, portfolio credit risk models need to be updated periodically as new data arrive and/or old data are revised. With one new month of data, this means that the final date index N is increased to $N + 1$. For this monthly real-time updating procedure, we always apply re-initialization, where the relevant means and variances–covariances used to generate the initial particle cloud are computed with the updated SMC parameter particle set from the previous run up to time N. Then one can apply the same recursive procedure, as described in Subsection 1.3.3. Furthermore, one can update all self-normalized statistics shown in Subsection 1.3.4 to reflect the additional one more pseudo-posterior means to the sequence.

As for this technical report, the initial parameter estimation by SMC is carried out for all calibration groups on June 11 2018 using the data up to the end of May 2018. Additional implementation details on the calibration are given in Section 3.

2 Input Variables and Data

Subsection 2.1 describes the input variables used in the quantitative model. In principle, the same set of input variables is common to most of the economies under the CRI's coverage. Going further, the CRI system starts to identify different input variables specific to different economies (e.g., China and India). The effect of each of the variables on the PD output will be discussed in the empirical analysis of Section 4.

Subsection 2.2 gives the data sources and relevant details of the data sources. There are two categories of data sources: current and historical. Data sources used for current data need to be updated in a timely manner so that daily updates of PD forecasts are meaningful. They also need to be comprehensive in their current coverage of firms. Data sources that are comprehensive for current data may not necessarily have comprehensive historical coverage for different economies. Thus, other data sources are merged in order to obtain comprehensive coverage of historical and current data.

Subsection 2.3 indicates the fields from the data sources that are used to construct the input variables. For some of the fields, proxies need to be used for a firm if the preferred field is not available for that firm.

Subsection 2.4 discusses the definition and sources of defaults and of other exits used in the CRI.

2.1 Input Variables

Following the notation that was introduced in Section 1, firm i's input variables at time $t = n\Delta t$ are represented by the vector $X_i(n) = (W(n), U_i(n))$ consisting of a vector $W(n)$ that is common to all firms in the same economy, and a firm-specific vector $U_i(n)$ which is observable from the date the firm's first FS is released, until the month end before the month in which the firm exits, if it does exit.

In Duan et al. [2012], different variables that are commonly used in the literature were tested as candidates for the elements of $W(n)$ and $U_i(n)$: the 2 common variables and 10 firm-specific variables were selected as having the greatest predictive power for corporate defaults in the United States. In the current stage of development, the set of 16 covariates beyond the past 12 variables, as described below, is generally used for all economies but China. In an ongoing effort, future development will include variable selection for firms in different economies.

- Common variables

  The vector $W(n)$ contains four elements, which are:

  1. Stock index return: the trailing one-year simple return on a major stock index of the economy;
  2. Interest rate: a representative 3-month short-term interest rate standardized from the data available point until now;
  3. Financial Aggregate DTD: median DTD of financial firms in each economy/country inclusive of those foreign financial firms whose primary shock exchange is in this economy/country;
  4. Non–financial Aggregate DTD: median DTD of non–financial firms in each economy/country inclusive of those foreign financial firms whose primary shock exchange is in this economy/country.

Stock index return incorporates the following two treatments. First, we use unified currencies for 6 groups of economies: China (CNY), India (INR), Asia-Pacific Developed (USD), Emerging Market (USD), Europe (EUR), and North America (USD). Second, we winsorize the unified return over the range of $[5\%, 95\%]$ for 3 groups of economies: Asia-Pacific Developed, Emerging Market, and Europe.

Interest rate is standardized in the way of demeaning each series and then scaling the demeaned values so that the standard deviation equals one, except for China and India. The treatment specific to the Eurozone is detailed in Subsection 3.3.

Each of the aggregate DTDs is only applicable to firms in the corresponding category. In short, the number of covariates used for default prediction is 16 including 12 firm–specific variables, as will be discussed below. China, however, differs from other economies/countries where the two aggregate DTDs are not applicable, because they offer no informational value above and beyond what have already been captured. The number of covariates for China is thus still 14.

- Firm-specific variables

  The 12 firm-specific input variables are transformations of measures of 6 different firm characteristics. The 6 firm characteristics are:

  1. volatility-adjusted leverage;
  2. liquidity;
  3. profitability;
  4. relative size;
  5. market mis-valuation/future growth opportunities; and
  6. idiosyncratic volatility.

Volatility-adjusted leverage is measured as the DTD in a Merton-type model. The calculation of DTD used by the CRI allows a meaningful DTD for financial firms, a critical sector that must be excluded from most DTD computations. This calculation is detailed in Section 3.

Liquidity is measured as a log ratio of cash and short-term investments to total assets for financial firms and a log ratio of current assets to current liabilities for non–financial firms. Profitability is measured as a ratio of net income to total assets. Relative size is measured as the logarithm of the ratio of market capitalization to the economy's median market capitalization.

Duan et al. [2012] transformed these first four characteristics into level and trend versions of the measures. For each of these characteristics, the level is computed as the one-year average of the measure, and the trend is computed as the current value of the measure minus the one-year average of the measure. The level and trend of a measure have seldom been used in the academic or industry literature for default prediction, and Duan et al. [2012] found that using the level and trend significantly improves the predictive power of the model for short-term horizons.

To understand the intuition behind using level and trend of a measure as opposed to using just the current value, consider the case of two firms with the same current value for all measures. If the level and trend transformations were not performed, only the current values would be used and the two firms would have identical PD. Suppose that for the first firm the DTD had reached its current level from a high level, and for the second firm the DTD had reached its current level from a lower level (see Fig. 2). The first firm's leverage is increasing (worsening) and the second firm's leverage is decreasing (improving). If there is a momentum effect in DTD, then firm 1 should have a higher PD than firm 2.

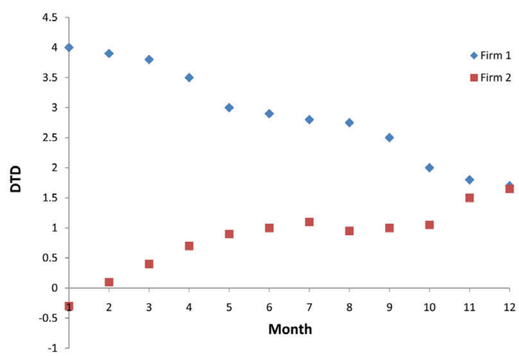

Figure 2: Two firms with all current values equal to each other, but DTD trending in the opposite direction.

Duan et al. [2012] found evidence of the momentum effect in DTD, liquidity, profitability and size. For the other two firm characteristics, applying the level and trend transformation did not improve the predictive power of the model.

As of this technical report, we further conduct additional treatments on liquidity and size. First, the level and trend of liquidity are respectively allowed to be sector–specific: financial firms, and non–financial firms. For financial firms, we take natural logarithm on the existing liquidity definition: log[(Cash + Short–term investments) / Total assets]. For non–financial firms, we refine liquidity as log(Current assets / Current liabilities) with the two current items in their financial statements. Second, size is redefined through the unified currency discussed above and then divided by the economy's median market capitalization over the past one year.

One of the remaining two firm characteristics is the market mis-valuation/future growth opportunities characteristic. This measure is taken as the "relative" market–to–book asset ratio (M/B) in the way of Individual firm's M/B divided by Economy M/B median at the same day that the individual M/B is calculated. In the CRI implementation, market–to–book asset ratio (M/B) is measured as a ratio of market capitalization and total liabilities to total assets. One can see whether the market mis-valuation effect or the future growth opportunities effect dominates this measure by looking at whether the parameter for this variable is positive or negative. This will be further discussed in the empirical analysis of Section 4.

The last firm characteristic is the idiosyncratic volatility which is taken as SIGMA, following Shumway [2001]. SIGMA is computed by regressing the daily returns of the firm's market capitalization against the daily returns of the economy's stock index, for the previous 250 days. SIGMA is defined to be the standard deviation of the residuals of this regression. Using daily returns is to ensure that SIGMA provides an accurate and timely measure of idiosyncratic risk of individual companies. Shumway [2001] reasons that SIGMA should be logically related to bankruptcy since firms with more variable cash flows and therefore more variable stock returns relative to a market index are likely to have a higher probability of bankruptcy.

Finally, the vector $U_i(n)$ contains 12 elements, consisting of:

1. Level of DTD.
2. Trend of DTD.
3. Level of log[(Cash + Short–term investments) / Total assets] for financial firms, abbreviated as CASH/TA.
4. Trend of CASH/TA for financial firms.
5. Level of log(Current assets / Current liabilities) for non–financial firms, abbreviated as CA/CL.
6. Trend of CA/CL for non–financial firms
7. Level of Net income / Total assets, abbreviated as NI/TA.
8. Trend of NI/TA.
9. Level of log (Firm market capitalization / Economy's median market capitalization over the past one year), abbreviated as SIZE.
10. Trend of SIZE.
11. Current value of Relative M/B defined as Individual firm's M/B divided by Economy M/B median, abbreviated as M/B.
12. Current value of SIGMA.

Note that every firm should belong to either a financial sector or a non–financial sector. Natually, this classification determines which the liquidity ratio between CASH/TA and CA/CL is used. When it comes to one financial firm, for example, we cannot use CA/CL level and trend among the 12 elements. Therefore, default prediction of each firm should depend on the rest of the 10 firm–specific variables. The data fields that are needed to compute DTD and short-term investments are described in Subsection 2.3. The remaining data fields required are straightforward and standard. The computation for DTD is explained in Section 3.

2.2 Data Sources

There are two data sources that are used for the daily PD forecast updates: Thomson Reuters Datastream and the Bloomberg Data License Back Office Product. Many of the common factors such as short-term interest rates and macroeconomic data are retrieved from Datastream.

Firm-specific data come from Bloomberg's Back Office Product which delivers daily update files by region via FTP after respective market closes. All relevant data is extracted from the FTP files and uploaded into the CRI database for storage. From this, the necessary fields are extracted and joined with previous months of data.

The Back Office Product includes daily market capitalization data based on closing share prices and also includes new FSes as companies release them. Firms will often have multiple versions of FSes within the same period, with different accounting standards, filing statuses (most recent, preliminary, original, reclassified or restated), currencies or consolidated/unconsolidated indicators. A major challenge lies in prioritizing these FSes to decide which data should be used. The priority rules are described in section 3.

The firm coverage of the Back Office Product is of sufficient quality that over 34,000 firms can be updated on a daily basis in the 128 economies under the CRI's coverage. While the current coverage is quite comprehensive, historical data from the Back Office Product can be sparse for certain economies. For this reason, various other databases are merged in order to fill out the historical data. The other databases used for historical data are: a database from the Taiwan Economics Journal (TEJ) for Taiwanese firms; a database provided by Korea University for South Korean firms; data from Prowess for Indian firms; and the Compustat for United States.

With all of the databases merged together and for the 128 economies under CRI's coverage, around 68,000 exchange-listed firms are in the CRI database. The historical coverage of the firm data goes back to the early 1990s. In order to be included in our coverage, a company needs to have common equity traded on a stock exchange. Of these 128 economies, 88 economies inclusive of Qatar as a new economy have their own stock exchange (see Table A.2). For the other 40 economies under the CRI coverage, we cover companies domiciled in the economy that are quoted on a foreign exchange, either because those economies do not have a stock exchange or because data issues are preventing us from including the companies listed on the local exchange (see Table A.3). For these reasons, we exclude the following two economies for the CRI products calibrated in June 11 2018 (May calibration): Dominican Republic and Niger Republic.

2.3 Constructing Input Variables

The chosen stock indices and short-term interest rates for the 88 economies with their own stock exchanges under the CRI's current coverage are listed in Tables A.5 and A.6, respectively. All economies are listed by their three letter ISO code given in Table A.4.

Most of the firm-specific variables can be readily constructed from standard fields from firms' FSes in addition to daily market capitalization values. The only two exceptions are the DTD and the liquidity measure.

The calculation for DTD is explained in section 3. In the calculation, several variables are required. One variable is a proxy for a one-year risk-free interest rate, and the choices for each of the 88 economies are listed in Table A.7. Total assets, long-term borrowing and total liabilities are also required, but can be obtained from standard FS fields easily.

Total current liabilities are also required, and due to the relatively large numbers of firms that are missing this value, proxies have to be found. The preferred Bloomberg field for this is BS_CUR_LIAB. If this is missing, then the sum of BS_ST_BORROW, BS_OTHER_ST_LIAB, BS_CUST_ACCPT_LIAB_CUSTDY_SEC (customers' acceptance and liabilities/custody securities) and BS_SEC_SOLD_REPO_AGRMNT is used. If one, two or three of these are missing, zero is inserted into those fields, but at least one of the four fields is required.

The liquidity measure requires different fields for financial and non-financial firms. For non–financial firms, the two elements of "CA/CL" come from BS_CUR_ASSET_REPORT and BS_CUR_LIAB, respectively: log(Current assets / Current liabilities). For financial firms, the numerator of "CASH/TA", (Cash + Short-term investments), is taken as the sum of BS_CASH_NEAR_CASH_ITEM, ARD_SEC_PURC_UNDER_AGR_TO_RESELL (securities purchased under agreement to re-sell), ARD_ST_INVEST, and BS_INTERBANK_ASSET. If one or two of the last three fields are missing, zero is inserted for those fields, but at least one field is required.

The "ARD" prefix indicates that these are "as reported" numbers directly from the FSes. As such, for some firms these fields may need to be adjusted to the same units before adding them to other fields.

To summarize, the firm-specific variables include: DTD, Cash/TA, CA/CL, NI/TA, SIZE, M/B, and SIGMA, and the statistics grouped by economy are listed in Table A.8.

2.4 Data for Corporate Events

The CRI database contains 11,604 default events and 57,002 other exits events from 1990 to the end of May 2018. The corporate events come from numerous sources, including Bloomberg, Compustat, CRSP, Moodys reports, TEJ, exchange websites and news sources. Moreover, in order to enhance default coverage, from December 2015, the CRI team has started to use "defaults" reported by major credit rating agencies as an additional data source.

The default events that are recognized by the CRI can be classified under one of the following events:

(1) Bankruptcy filing, receivership, administration, liquidation or any other legal impasse to the timely settlement of interest and/or principal payments;

(2) A missed or delayed payment of interest and/or principal, excluding delayed payments made within a grace period;

(3) Debt restructuring/distressed exchange, in which debt holders are offered a new security or package of securities that result in a diminished financial obligation (e.g., a conversion of debt to equity, debt with lower coupon or par amount, debt with lower seniority, debt with longer maturity).

The more precise sub-categories of default corporate actions are listed in Table A.9.

Delisting due to other reasons such as failure to meet listing requirements, inactive stock prices or M&A are counted as "other exits" and are not considered as default. Especially, if a firm has stale stock price for more than a year but has no record on experiencing any credit events, we will assume that it has been suspended and exited from its stock exchange. If two credit events of the same type happen in a row or a default event happens followed by another event of either type, we only keep the first event assuming that the series of events arise from the same source of financial distress. However, if firms are delisted from an exchange and then experience a default event within 365 calendar days of the delisting, we will only keep the default event, and any information between the two dates won't be used. Technical defaults such as covenant violations are not included in our definition of default. The exit events that are not considered as defaults in the CRI system are listed in Table A.10.

In addition to the aforementioned events, there are still cases that require special attention and will be assessed on a case-by-case basis, e.g., subsidiary default. As a general rule, the CRI does not consider related party-default (e.g., subsidiary bankruptcy) as a default event. However, when a non-operating holding parent company relies heavily on its subsidiary, bankruptcy by the subsidiary will cause a considerable economic impact on the parent company. Such cases will be reviewed, and final classifications will be made.

Complete statistics of the total number of firms, number of defaults, and number of other exits in each of the 88 economies from 1990 to 2017 are listed in Table A.11.

3 Implementation Details

Section 1 described the modeling framework underlying the current implementation of the CRI system. It focused on theory rather than the details encountered in an operational implementation. The present section describes how the CRI system handles more specific issues.

Subsection 3.1 describes implementation details related to data, mainly dealing with data cleaning and missing data. Subsection 3.2 describes the specific computation of DTD used by the CRI system that leads to meaningful DTD for financial firms. Subsection 3.3 explains how the calibration previously described in Subsection 1.2 can be implemented. Subsection 3.4 gives the implementation details relevant to the daily output. This includes an explanation of the various modifications needed to compute daily PDs so that the daily PDs are consistent with the usual month end PD and a description of the computation of the aggregate PDs provided by the CRI.

3.1 Data Treatment for Calibration

Fitting data to monthly frequency: Historical end of month data for every firm in an economy is required to calibrate the model. For daily data such as market capitalization, interest rates and stock index values, the last day of the month for which there is valid data is used.

Up to the October 2012 calibration, FS variables data were used, starting from the period end of the statement lagged by 3 months. This is to ensure that predictions are made based on information that was available at the time the prediction was made. However, this treatment can be over-conservative, and many companies actually release their FSes quicker than 3 months. Therefore, we implement a new logic, and we start using the values in an FS as soon as its latest revision was put into the CRI database, unless the FS' release was delayed for more than 3 months. If there was no revision to an FS, the originally released FS is used. Whenever the latest revision is available more than 3 months after the period end, we revert to the previous logic. We start including the FS before the latest revision is actually available as a compromise, to avoid situations like later minor revisions of the FS holding back more up-to-date information. It should be noted that the new approach was only applied for FS input into the CRI database after February 2011, as the revision dates were not accurately recorded before this date. The CRI considers FS variables to be valid for one year without restriction, after they were first used.

Priority of FSes with the same period end: As described in Subsection 2.2, data provided in Bloomberg's Back Office Product can include numerous versions of FSes within the same period. If there are multiple FSes with the same period end, priority rules must be followed in order to determine which to use. The formulation and implementation of these rules are major challenges and areas of continuing development.

The first rule is to prioritize by consolidated/unconsolidated status. This rule applies to all economies, however, special treatment is imposed on firms in the "diversified financial services" sector in South Korea and Taiwan. In this sector of the two economies, firms issue unconsolidated FSes more frequently than consolidated ones. As a result, this prioritization rule can lead to cases where the FSes chosen switch between unconsolidated and consolidated ones on a regular basis. In South Korea and Taiwan, where corporate structures are biased toward large holding companies, this switching may substantially distort the DTD calculation for these holding companies. Therefore, as of October 2013 calibration, in the case of South Korea, and November 2013 calibration, in the case of Taiwan, if a company has released at least one consolidated FS over the last 12 months, all unconsolidated FS will be ignored.

If, after the first prioritization rule has been applied, there are still multiple FSes, the second rule is applied. This is prioritization by fiscal period. In most economies, annual statements are required to be audited, whereas other fiscal periods are not necessarily audited. The order of priority from highest to lowest is, therefore: annual, semi-annual, quarterly, cumulative, and finally other fiscal periods. We have observed that the capital structure breakdown reported by Australian domiciled-banks differs between annual and semi-annual reports, leading to DTD calculations that are not meaningful. Because of this, as of October 2013 calibration, we only use data from annual FSes for Australian banks.

The third prioritization rule is based on filing status. The "Most Recent" statement is used before the "Original" statement, which is used before the "Preliminary" statement.

The final prioritization rule is based on the accounting standard. As more and more countries adopt the International Financial Reporting Standards (IFRS) as their mandatory accounting standard, FSes that are reported using IFRS are given higher priority than they were before. The revised rule is implemented from the 2014 October calibration and is described as follows. For the countries with mandatory IFRS adoption, FSes under IFRS are now given the highest priority after their respective mandatory adoption dates. Before the mandatory adoption dates and for countries without mandatory IFRS adoption, FSes under the Generally Accepted Accounting Principles (GAAP) have the highest priority. If an FS does not indicate its accounting standard, it will not be used.

Having all the prioritization descriptors in place, we rank all the FSes available in the database from the highest priority to the lowest. If there are FSes where all the financial information needed in our model is present, the FS with the highest ranking will be chosen. If instead there is no such FS, we will pick the values variable by variable. For example, the total liability is taken from the highest ranked FS with this information available, while the total asset can be from another FS, which ranks the highest among those bearing this information and having the same FS period end. This treatment is to get as much information as possible and to accommodate the fact that Bloomberg occasionally only revises the variables that have changed values, leaving the other fields NaN.

One variable that requires special attention is the net income. Net income is a flow variable and needs to be adjusted based on the fiscal period of the FS. More specifically, we transform the net income into a monthly net income by dividing the net income by the number of months that the FS covers. For example, the monthly net income can be computed from the annual net income divided by 12, the semi-annual net income divided by 6 and the quarterly net income divided by 3. When the monthly net income can be obtained from different sources simultaneously, the quarterly net income will have the highest priority (followed by the cumulative quarterly, semiannual, annual, and others) because it covers a more recent period of time.

Treatment of stale market capitalization prices: The market capitalization of a firm is required in a few input variables: DTD, SIZE, M/B, and SIGMA. For most firms, the market capitalization is available from Bloomberg on a daily basis.

A check on the trading volume of shares is used to remove stale prices. Specifically, if there are more than two consecutive days of identical market capitalization prices, subsequent identical prices are removed only if the trading volume is equal to zero. This is to avoid, for example, cases where the shares of a company are under a trading suspension but the market capitalization data is incorrectly carried forward.

An exception is for Indian companies, where it is common for some companies to have market capitalizations reported only once a month with several consecutive months having identical prices and positive trading volume. These prices are very likely not to be accurate reflections of the firms' value. So, the trading volume is not checked for Indian firms and market capitalizations are excluded after more than two repeated prices.

For some firms, the market capitalization data is not available for some periods. To fill in the blanks, we use the shares outstanding obtained from the previously available market capitalization divided by the price on that day as a proxy. If the market capitalization data is missing for more than a year, we use the share price multiplied by the shares outstanding listed on the balance sheet and then multiplied again by the adjustment factor that Bloomberg provides to account for splits, dividends, etc. If there is still market capitalization missing in the data, then shares outstanding from other data sources including Compustat and Korean University Database are used.

Currency conversion: Currency conversions are required if the market capitalization or any of the FS variables are reported in a currency different than the currency of the economy. If a currency conversion is required, the foreign exchange rate used is the one reported at the relevant market close. For firms traded in most of the Asian economies and Asia-Pacific, the Tokyo closing rate is used; for firms traded in Europe, Africa, and Middle East, the London closing rate is used; and for firms traded in North and Latin America, the New York closing rate is used. For market capitalizations, the FX rate used is for the date that the market capitalization is reported. For FS variables, the FX rate used is for the date of the period end of the statement.

As of December 2017, we proceed with the unified currency treatment about stock index return for each calibration group of economies: China (CNY), India (INR), Asia-Pacific Developed (USD), Emerging Market (USD), Europe (EUR), and North America (USD). This attempt is made to prevent currency distortion in assessing default prediction. Similarly, we apply the currency adjustment to market capitalization, total liabilities, and total assets, all of which are used to compute the M/B ratio.

Treatment for mergers and acquisitions (M&A): M&A events are common occurrences in the economic world. For our purpose, we define the M&A events as the cases where a firm ("acquirer") acquires partial or full ownership of another firm ("target"). Once an M&A deal is completed, the market capitalization of the acquirer changes immediately, reflecting the restructure of the acquirer. However, its FSes do not usually immediately reflect the new situation due to the fact that they are only released on a periodic basis. As a result, the DTD and market-to-book ratio, which are important inputs for the PD computation, will be distorted due to a mismatch in the market capitalization and the FS variables. In order to ensure the accuracy and reliability of our PD estimates, some special treatments are taken for PD calculations to companies whose financials are presumably significantly affected by the M&A events. The treatments are only applied to the acquirers.

The treatment starts with the screening of the important M&A deals. Only the important M&A deals are treated, assuming that the unimportant ones would not significantly affect a firm's corporate structure. An M&A deal is considered important if it satisfies the following three criteria :

1. Upon the deal's completion, the acquirer owns 20% or more of the target company.
2. The size of the deal is material to the acquirer. This is measured in terms of total assets. If α is the percentage of the target that is being acquired, the size is considered material if the product of α and the total assets of the target is greater than or equal to 20% of the total assets of the acquirer.
3. The change in market capitalization is material, with the largest absolute daily market capitalization return, within 20 days of the M&A completion day, larger than or equal to 5%.

One thing to note in implementation is that some targets stopped producing financial statements years before the M&A events. As a result, they may not have a valid value of total asset (needed for testing criterion 2) on the deal completion date. In this case, we use their last available value within 2 years before the deal completion as a substitute. If the last available value is beyond the 2-year range, we think that the data is not informative enough to reflect the financial situation upon deal completion and thus skip this particular case.

In order to mitigate the mismatching problem between the market capitalization and FS variables, we make the simplest and most conservative treatments, which are in line with the fundamental accounting standards. The treatment period will begin on the deal completion date and end when the first financial statement that reflects the post-M&A situation becomes available, which varies across economies and can range from 3 months to a few years. After identifying the important M&A deals, which must have had an ownership level of equal or more than 20%, we treat them in two different ways:

1. If the acquirer owns 20-50% (excluding 50%) of the target upon deal completion, the "Equity Method" is used to treat the financial statement variables. Under the "Equity Method", the total asset of the acquirer will increase by a proportion, which is the percentage of ownership acquired in this deal, of the targets equity. Its net income will increase by the same proportion of the target's net income. In contrast, other financial statement variables will stay the same.
2. If the acquirer owns 50-100% (including 50%) of the target upon deal completion, the "Acquisition Method" is used to adjust the financial statement variables. By using this method, we assume that the financial manager of the acquirer consolidates the financial statements of both entities. As a consequence, the financial statement variables, including total liability, total asset, and cash and marketable securities, take the simple sum of the values from both entities. The net income will still increase by a proportion (the percentage of ownership acquired in this deal) of the targets net income, simply because it is the profit attributed to the shareholders.

After constructing the hypothetical financial statement data in the above-mentioned way, we use them to compute the DTD and the historical monthly PDs wherever applicable. Note that we do not let the hypothetical values enter the model's calibration process. With enough data points in the database to robustly calibrate the model parameters at the economy or region level, we can afford to disregard a small portion of data for the M&A period during which we believe them to be mismatched. After getting the model parameters, however, we not only use the hypothetical values to re-calibrate the firm-specific DTD parameters and re-calculate the DTD values, we also use them to adjust other variables with financial information. This is to guarantee that the PDs during the treatment period are properly calculated.

Treatment for missing values and outliers: Missing values and outliers are dealt with by a three-step procedure. In the first step, the 10 firm-specific input variables are computed for all firms and all months. In this step, the extreme values will be calculated, and the missing values will be determined. In the second step, outliers are eliminated by winsorization. In the final step, missing values are replaced under certain conditions.

The first step is to compute the input variables and to determine which are missing. As mentioned previously, FS variables are carried forward for one year after the date that they are first used. The date that they are first used is generally three months after the period end of the statement. If no FS is available for the company within this year, then the FS variable will be missing. For market capitalization, if there is no valid market capitalization value within the calendar month, then the value is set to missing.

With regard to the level variables, their values in the current and the last 11 months are averaged to compute the level. A minimum of 6 observations in the 12-month range are required to calculate the level variables. If fewer than 6 observations exist in this case, the level variables will bear missing values. However, this condition is not enforced during the initial 6 months after the firm releases the first financial statement.

To compute the trend variables, the level is subtracted from the current month value. If the current month value is missing, the trend variable is set to be the last valid value during the previous one year.

The value of M/B is set to be missing if any of the following values are missing: market capitalization, total liabilities, or total assets of a firm. For the computation of SIGMA, at least 50 valid returns over the last 250 days of possible returns are required for the regression. If there are less than 50 valid returns, SIGMA is set to be missing.

In this way, the 8 trend and level variables as well as M/B and SIGMA are computed and identified as missing or present. Winsorization can then be performed as a second step to eliminate outliers. The volume of outliers is too large to be able to determine whether each one is valid or not, so winsorization applies a floor and a cap on each of the variables. The historical 0.1 percentile and 99.9 percentile for all firms in the economy are recorded for each of the 10 variables. Any values that exceed these levels are set to equal these boundary values.

With a winsorization level of 0.1 and 99.9 percentile, the boundary values still may not be reasonable. For example, NI/TA levels of nearly -25, meaning an annual net income -25 times larger than the total assets of a firm, has been observed at this stage. In these cases, a more aggressive winsorization level is applied, until the boundary values are reasonable. Thus, the winsorization level is economy- and variable-specific, and will depend on the data quality for that economy and variable. Winsorization levels different from the default of 0.1 percentile and 99.9 percentile are indicated in Table A.8. As for log variables log(x) such as CASH/TA and CA/CL, we should check first whether x is well defined with positive values. Otherwise, we assign the upper and lower bounds of the economy- and variable-specific winsorization level to these firms.

In addition to the special winsorization of the firm–specific variables, we also implement a winsorization of 5 and 95 percentiles for stock index return used as one of the common variables to the 3 groups of economies: Asia-Pacific Developed, Emerging Market, and Europe.

A third and final step can be taken to deal with missing values. If during a particular month, no variable is missing for a particular firm, the PD can then be computed. If 6 or more of these 10 variables are missing, there is deemed to be too many missing observations and no replacement shall be made.

If between 1 and 5 variables are missing out of the 10, the first step is to trace back for at most 12 months to use previous values of these variables instead. If this does not succeed in replacing all of the variables, a replacement by sector medians is done. A firm's sector during a certain month is classified as either financial or non-financial, which is based on its Bloomberg industry sector code during that month. As of January 2015, the sector median replacement is no longer implemented in the calibration process but still in the PD computation. One special case is that the sector replacement is not done if it results in a relative change in the historical PD of 10% or more when the initial PD was at or above 100 bps, or an absolute change in the historical PD of 10 bps or more when the initial PD was below 100 bps.

One thing to note is that in the initial phase of a company - 6 months or even longer after its IPO - the data availability and quality are relatively low due to, for example, the delay in the issuance of FSes or illiquid trading. As observed in our data, replacing the missing values during this period with a sector median sometimes results in extreme spikes and falls in the company's PD. These extreme values are not easily detected, because in the beginning of a company's history, there are not many previous PD values to compare to as can be done later in the company's history. In order to avoid this, as of the 2015 January calibration, we set the rule to start treating the missing values only from the month when both the DTD level and trend are available and finite. By doing so, we make the PDs in the beginning of a company's history more reflective of its true credit quality.

Inclusion/exclusion of companies for calibration: Firms are included within an economy for calibration when the primary listing of the firm is on an exchange in the economy. This ensures that all firms within the economy are subject to the same disclosure and accounting rules. There are a relatively small number of firms that are listed in multiple economies. For example, Bank of China Ltd is listed both in Hong Kong Stock Exchange and China's Shanghai Stock Exchange. Based on Bloomberg's classification of its primary listing, Bank of China Ltd is assigned to the calibration group of Asia-Pacific rather than China.

In the US, firms traded on the OTC markets or the Pink Sheets are not considered as exchange listed so are not included in calibration or in the reporting of PD forecasts. Many of these firms are small or start-up firms. Including this large group of companies would skew the calibration and the aggregate results. The TSX Venture Exchange in Canada also contains only small and start-up firms, so firms listed here are also excluded.

Other exclusions include Taiwan's Taipei Exchange, Vietnam's Hanoi UPCoM, Switzerland's OTC-X BEKB, Brazil's Soma and Romania's RASDAQ. To identify the smaller markets outside of the US and Canada is challenging due to data availability. However, continuing work is being done in the CRI system to exclude firms that are not listed on major exchanges within a country.

3.2 Distance-to-Default Computation

The DTD computation used in the CRI system is not a standard one. Standard computations exclude financial firms, which is of course a critical part of any economy. Thus, the standard DTD computation must be extended to give meaningful estimates for financial firms as well. Duan and Wang [2012] have provided a review of different DTD calculations with several examples for financial and non-financial firms.

The description of the specialized DTD computation starts with a brief description of the Merton [1974] model. Merton's model makes the simplifying assumption that firms are financed by equity and a single zero-coupon bond with maturity date T and principal L. The asset value of the firm V_t follows a geometric Brownian motion:

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}B_t. \tag{30}$$

Here, B_t is the standard Brownian motion, μ is the drift of the asset value in the physical measure, and σ is the volatility of the asset value. Following the Merton [1974] model, the probability of the company's default at time T evaluated at time t is $\Pr_t(V_T \leq L)$, from Eq. (30), we can derive $\Pr_t(V_T \leq L) = N(-\mathrm{DTD}_t)$, where DTD at time t is defined as:

$$\mathrm{DTD}_t = \frac{\log\left(\frac{V_t}{L}\right) + \left(\mu - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}. \tag{31}$$

The standard KMV assumptions given in Crosbie and Bohn [2003] are to set the time to maturity $T-t$ at a value of one year, and the principal of the zero-coupon bond L to a value equal to the firm's current liabilities plus one half of its long-term debt. Here, the current liabilities and long-term debt are taken from the firm's FSes. If the firm is missing the current liabilities field, then various substitutes for this field can be used, as described in Subsection 2.3.

This is a poor assumption of the debt level for financial firms, since they typically have large liabilities, such as deposit accounts, that are neither classified as current liabilities nor long-term debt. Thus, using these standard assumptions means ignoring a large part of the debt of financial firms.

To properly account for the debt of financial firms, Duan [2010] included a fraction δ of a firm's other liabilities. The other liabilities are defined as the firm's total liabilities minus both the short and long-term debt. The debt level L then becomes the current liabilities plus half of the long-term debt plus the fraction δ multiplied by the other liabilities, so that the debt level is a function of δ. The standard KMV assumptions are then a special case where $\delta = 0$.

The fraction δ can be optimized along with μ and σ in the transformed-data maximum likelihood estimation method developed in Duan [1994, 2000]. As asset value is unobservable, it has to be implied from market equity value. Note that equity holders receive the excess value of the firm above the principal of the zero-coupon bond and have limited liability, so the equity value at maturity is: $\max(V_T - L, 0)$. This is just a call option payoff on the asset value with a strike value of L. Thus, the Black-Scholes option pricing formula can be used to calculate the equity value at times t before T,

$$E_t = V_t N(d_+) - e^{-r(T-t)} L N(d_-), \tag{32}$$

where r is the risk-free rate, $N(\cdot)$ is the standard normal cumulative distribution function,

$$d_{\pm} = \frac{\log\left(\frac{V_t}{L}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{33}$$

and $L \equiv L(\delta) = \text{Current Liabilities} + 1/2 \cdot \text{Long-term Debt} + \delta \cdot \text{Other Liabilities}$ as mentioned before. Then we can express the likelihood function of the observed equity values by viewing the equity values as the transformed data from pricing formula in Eq. (32). It should be noted that the transformation involves the unknown asset volatility. By standard transformation theory, the likelihood of observed equity values must equal the product of the likelihood of the asset values (implied by equity values) and the Jacobian of the inverse transformation (from the equity value back to the asset value). Moreover, following Duan et al. [2012], the firm's market value of assets is standardized by its book value A_t, so that the scaling effect from a major investment or financing by the firm will not distort the time series from which the parameter values are estimated. Thus, the log-likelihood function based on equity prices is:

$$\begin{aligned}\mathcal{L}(\mu,\sigma,\delta) = & -\frac{n-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\log(\sigma^2 h_t) - \sum_{t=2}^{n}\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\right) \\ & - \sum_{t=2}^{n}\log[N(\hat{d}_+(\hat{V}_t(\sigma,\delta),\sigma,\delta))] \\ & - \frac{1}{2\sigma^2}\sum_{t=2}^{n}\frac{1}{h_t}\left[\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\times\frac{A_{t-1}}{\hat{V}_{t-1}(\sigma,\delta)}\right) - \left(\mu - \frac{\sigma^2}{2}\right)h_t\right]^2, \end{aligned} \tag{34}$$

where n is the number of days with observations of the equity value in the sample, $\hat{V}_t$ is the implied asset value found by solving Eq. (32), $\hat{d}_+$ is computed with Eq. (33) using the implied asset value, and h_t is the number of trading days as a fraction of the year between observations $t-1$ and t. Notice that the implied asset value and $\hat{d}_+$ are dependent on δ by virtue of the dependence of L on δ.

Implementation of DTD computation: The DTD at the end of each month is needed for every firm in order to calibrate the forward intensity model. A moving window, consisting of the last one year of data before each month end is used to compute the month end DTD. Daily market capitalization data based on closing prices is used for the equity value in the implied asset value computation of Eq. (32). If there are fewer than 50 days of valid observations for the DTD input variables (market capitalization, FS variables, and interest rate), the DTD value is set to be missing. An observation is valid if there is positive trading volume that day. If the trading volume is not available, the observation is assumed to be valid if the value for the market capitalization changes often enough. The precise criterion is as follows: if the market capitalization does not change for three days or more in a row, the first day is taken as a valid observation, and the remaining days with the same value are set to be missing.

A straightforward idea for the DTD computation is to first estimate the three variables μ, σ and δ via maximizing the log-likelihood function (34) over $\sigma \geq 0$ and $0 \leq \delta \leq 1$, and then to calculate the DTD from Eq. (31). Let $(\hat{\mu},\hat{\sigma},\hat{\delta})$ be an optimal solution to the maximization problem. By direct calculation, it is not hard to see that

$$\hat{\mu} = \frac{\hat{\sigma}^2}{2} + \frac{1}{\sum_{t=2}^{n} h_t}\log\left(\frac{\hat{V}_n(\hat{\sigma},\hat{\delta})}{A_n}\times\frac{A_1}{\hat{V}_1(\hat{\sigma},\hat{\delta})}\right). \tag{35}$$

In view of this, maximizing the three-dimensional function $\mathcal{L}(\mu,\sigma,\delta)$ can be equivalently reduced to maximizing the two-dimensional function $\tilde{\mathcal{L}}(\sigma,\delta)$ taking the form

$$\begin{aligned}\tilde{\mathcal{L}}(\sigma,\delta) = &-\frac{n-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\log(\sigma^2 h_t) - \sum_{t=2}^{n}\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\right) \\ &- \sum_{t=2}^{n}\log N(d_+) - \frac{1}{2\sigma^2}\left\{\sum_{t=2}^{n}\frac{1}{h_t}\times\left[\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\times\frac{A_{t-1}}{\hat{V}_{t-1}(\sigma,\delta)}\right)\right]^2\right. \\ &\left. - \frac{1}{\sum_{t=2}^{n}h_t}\left[\log\left(\frac{\hat{V}_n(\hat{\sigma},\hat{\delta})}{A_n}\times\frac{A_1}{\hat{V}_1(\hat{\sigma},\hat{\delta})}\right)\right]^2\right\}. \end{aligned} \tag{36}$$

However, with quarterly FSes there will never be more than three changes in the corporate structure (defined in this model by L and A_t) throughout the year, leading to possibly unstable estimates of δ. This problem is mitigated by performing a two-stage optimization for σ and δ.

In the first stage, the maximization of $\tilde{\mathcal{L}}(\sigma,\delta)$ for each firm is performed over both σ and δ. For each firm, at the first month in which DTD can be computed, the maximization is constrained in $\sigma \geq 0$ and $0 \leq \delta \leq 1$. Thereafter, at month n, the maximization is still constrained in $\sigma \geq 0$ while δ is constrained in the interval $[\max(0,\hat{\delta}_{n-1}-0.05), \min(1,\hat{\delta}_{n-1}+0.05)]$, where $\hat{\delta}_{n-1}$ is the estimate of δ made in the previous month. In other words, a 10% band around the previous estimate of δ (where that band is floored with 0 and capped with 1) is applied so that the estimates do not fluctuate too much from month to month.

However, for many firms, the estimate of δ would frequently lie on the boundary of the constraining interval, meaning that the estimates of δ were not stable. Therefore, a second stage is implemented to impose greater stability. Within the same calibration group, all firms in the same sector (Bloomberg 10-industry sectors classification) are assumed to share the same estimate of δ, chosen to be the average of all its individual estimates. However, for some small economies, especially in their early years, the average of δ is still observed to be not stable due to some sector or even the whole calibration group has only few individual estimates of δ. To well handle such cases, a threshold rule at each time of estimation is applied under the following conditions: a) If a sector has fewer than 10 individual estimates, the shared estimate of δ will be set to the average of whole calibration group instead of the sector average; b) furthermore, if the whole calibration group still has fewer than 10 individual estimates, the shared estimate of δ is deemed not available. Accordingly, with δ being fixed to be the sector average on the calibration group level, the original maximization of $\tilde{\mathcal{L}}(\sigma,\delta)$ is reduced to a one-dimensional maximization in σ for each firm.

Since the first stage is done to obtain a stable sector-average estimate of δ, the criteria used to include a firm-month is more strict. In the first stage, a two-year window of FS variables, market capitalization, and interest rate is used instead of one year, and a minimum of 250 days of valid observations of the DTD input variables are required instead of 50. If a firm has less than 250 days of valid observations within the last two years of a particular month end, δ will not be estimated for that firm and that month end.

It was found that after applying the two-stage procedure described above, the estimate of μ was frequently unstable and could lower the explanatory power of DTD. For example, suppose a firm has a large drop in its implied asset value in January 2011, so that the estimated μ is negative for the DTD calculation at the end of December 2011. If there is little change in the company in January 2012, then the drop in implied asset value in January 2011 is no longer within the observation window for the DTD calculation at the end of January 2012. There will be a large increase in the estimated μ, resulting in a substantial improvement of the DTD just because of the moving observation window. To avoid this problem, we now set μ to be equal to $\sigma^2/2$. So in calculating DTD, the second term in the numerator of Eq. (31) is eliminated.

In summary, the DTD for each firm is computed using the sector average within a calibration group for δ in that month, and the estimate of σ based on the last year of data for the firm.

Carrying out this two-stage procedure would take about 70 hours of computation time on a single PC, given the millions of firm months that are required. However, each of the stages is parallelizable. In the first stage, the DTD can be computed independently between firms. In the second stage, once the sector averages of the δ have been computed for each month, the DTD can again be computed independently between firms. In the current CRI system, by using the NUS' high-performance computing facility, the DTD computational time has been greatly reduced thanks to the application of parallel computing.

3.3 Calibration

Implementation: As shown in Section 1, the calibration of the forward intensity model involves multiple maximum pseudo-likelihood estimations, where the pseudo-likelihood functions are given in Eq. (13). The maximizations are on the logarithm of these expressions, and the default parameters' maximization is performed independently from the non-default exit parameters. Parameter estimates for the entire horizon up to five years for the default and non-default exits can be obtained directly from the NS function.

A few input variables have an unambiguous effect on a firm's probability of default. Increments of both the level and trend of DTD, CASH/TA, CA/CL, and NI/TA should indicate that a firm is becoming more creditworthy and should lead to a decreasing PD. For large and relatively clean data sets such as the US, an unconstrained optimization leads to parameter values which mostly have the expected sign. For each of the DTD level and trend, CASH/TA level and trend, CA/CL level and trend, and NI/TA level and trend, the default parameters at all horizons are negative. A negative default parameter at a horizon means that if the variable increases, the forward intensity will decrease (based on Eq. (6)), so that the conditional default probability at that horizon will decrease.

Grouping for economies: There are not enough defaults in some small economies and calibrations of these individual economies are not statistically meaningful. In order to ensure that there are enough defaults for calibration, the 88 economies are categorized into groups according to similarities in their stage of development and their geographic locations. Within these groups, the economies are combined and calibrated together.

As of January 2015, Canada and the US remain in the North America calibration group, and the developed economies of Asia-Pacific (Australia, Hong Kong, Japan, Singapore, South Korea, Taiwan and New Zealand) form another calibration group. China and India, the two major emerging economies of Asia-Pacific are each calibrated as individual groups. All the European countries covered by the CRI are in a single calibration group. The other emerging economies of Asia-Pacific, Latin America, Middle-East, and Africa form the "emerging markets" calibration group, which now includes 9 African economies: Botswana, Ghana, Kenya, Malawi, Mauritius, Namibia, Rwanda, Tanzania, and Uganda. Detailed grouping can be found in Table A.4.

All economies in the same calibration group share the same coefficients for all common variables except for the 3-month interest rate variable. In particular, we apply standardization to each economy's interest rate time series, except for China and India. First, we subtract the historical month-end mean from the 3-month interest rate variable in order to reflect the contemporary change relative to the historical average. We then scale the demeaned values so that the standard deviation equals one. Doing so allows to put all economies on the same scale so that the same interest rate parameter can be reasonably used on firms from different countries/economies.

We allow for a unique coefficient on the interest rate variable for each economy. However, certain treatments and exceptions apply due to various reasons. For New Zealand, it does not have enough default events to identify a separate coefficient. In this case, the actual interest rates are replaced with zeros throughout the whole time series. This is to disable the effect of interest rate in the particular calibration, but it will not induce bias based on the nature of the standardized interests. For the Eurozone economies, all of them use the standardized Germany's 3-month Bubill rate after the Eurozone was launched on January 1st, 1999. This aims to reflect more of the monetary rather than the sovereign credit conditions in those economies. Before joining the Eurozone, each of those economies except Germany uses own standardized interest rates, because none of them has enough default events before that date. Among the non-Eurozone economies, Denmark, Norway, Sweden, and UK have their own respective coefficients on the interest rate variable, but Iceland, Switzerland along with all the others share the same one. In the Emerging Markets group, only Indonesia, Malaysia, the Philippines, and Thailand have their own economy-specific coefficients on the interest rate variable. The Latin American subgroup has a universal coefficient for all the member economies, and all the others in the Emerging Markets group share the same coefficient.

One thing to note is that in addition to the unique coefficient on the interest rate variable, Indonesia also has its own coefficient for the relative size level as of October 2013.

Relative size: For the calibration data set, the median market cap of firms in an economy for each month end includes the market cap from the last trading day of each firm in the month. If a firm does not trade in a particular month, the firms market cap is not included in the median. For certain economies, many firms are illiquid and the median market cap experiences large variations due to the change in composition of firms rather than the market value of the firms. Another problem is data quality at the beginning of the historical sample: if a data provider starts including the market cap for a large number of firms in one month compared to the previous, there can be a large jump in the median market cap. Our research also reveals that debt–ridden countries (e.g., Venezuela) are usually susceptible to hyperinflation so that the market value of the firms under the severe economic turmoil is not trustworthy.

To avoid this problem, we use the economy's median market cap over the past one year as the divisor in the Relative Size variable:

1. We collect the whole market cap data of individual firms in a specific economy over the past one year.
2. We calculate the ratio of individual firm's market cap to the economy's median market cap calculated above.
3. We take a natural logarithm to reduce its variability.

3.4 Daily Output

Individual firms' PD: In computing the pseudo-log-likelihood functions in Eq. (13), only the end of month data is needed. The data needs to be extended to daily values in order to produce daily PDs.

For the level variables, the last 12 end-of-month observations (before averaging) are combined with the current value. The current value is scaled by a fraction equal to the current day of the month divided by the number of calendar days in the month. The earliest monthly value is scaled by one minus this fraction. The sum is then divided by the number of valid monthly observations, with the current value and the earliest monthly value jointly having the weight of one observation if either or both are not missing. Not performing this scaling can lead to an artificial jump in PD at the beginning of the month. When performing the scaling, the change in level is more gradual throughout the month.

SIGMA is computed by regressing the daily returns of the firm's market capitalization against the daily returns of the economy's stock index for the previous 250 days.

Aggregating PDs: The CRI provides term structures of the probability distributions for the number of defaults as well as the expected number of defaults for different groups of firms.

The companies are grouped by economy (using each firm's country of domicile), by sector (using the firm's Bloomberg industrial sector code) and sectors within economies.

To compute the probability distribution of the number of defaults, we use an algorithm which was originally reported in Anderson et al. [2003]. It assumes conditional independence and uses a fast recursive scheme to compute the necessary probability distribution. With the individual firms' PDs, the expected number of defaults is trivial to compute and is simply the sum of the individual PDs within each group. Note that while this algorithm is currently used to produce the probability distribution of the number of defaults within an economy or sector, it can easily be generalized to compute loss distributions for a portfolio manager, in which case the portfolio's exposure to each firm should be aggregated.

As of 8th July 2014, the display of the aggregate PDs on the RMI-CRI website started to adopt the simple median of the individual PDs within each group. This change will mitigate the effect from extreme outliers and synchronize with the aggregate display of the AS. It should be noted that the aggregate PDs using mean values are still accessible through the data downloading section on the website.

Inclusion of firms in aggregation: As explained in Subsection 3.1, firms are included in an economy for calibration if the firms' primary listing is on an exchange in that economy. This is to ensure that all firms in an economy are subject to the same disclosure and accounting requirements. In contrast, a firm is included in an economy's aggregate results if the firm is domiciled in that economy. This is because users typically associate firms with their economy of domicile rather than the economy where their primary listing is, if they are different. For example, the Bank of China has its primary listing in Hong Kong, but its economy of domicile is China so the Bank of China is included in the aggregation forecasts for China, and is included under China when searching for the individual PDs.

Treatment of companies after a default event: When a company experiences a default event, the CRI system discontinues the PD calculation for that company. However, if the company resumes operations after some time, it will be treated as a new company, and we continue to generate PD. The new company's PDs are not affected by the FS or market cap data prior to the event. So, the PDs calculated are independent of the PDs that were generated before the default event. On our website, the PDs are however displayed on a single graph for the convenience of our users.

4 Empirical Analysis

This section presents an empirical analysis of the CRI outputs for the 88 economies with their own exchange that are currently being covered. In Subsection 4.1, an overview is given of the default parameter estimates. Subsection 4.2 explains and provides the accuracy ratios for the different countries under the CRI coverage.

4.1 Parameter Estimates

With 60 months of forecast horizons, 17 variables (16 variables plus an intercept), and 6 different groups of economies, tables of the parameter estimates occupy over 20 pages and are not included in this Technical Report. In Figs. B.1 and B.2, the parameter estimates are from calibrations performed on June 11 2018 (May calibration) using data until the end of May 2018. As an example, plots of the default parameters for the US are given in Figs. B.1 and B.2 in Appendix B. In this part, a brief overview is given of the general traits and patterns seen in the default parameter estimations of the economies covered by the CRI.

Recall that if a default parameter for a variable at a particular horizon is estimated to be positive (negative) from the maximum pseudo-likelihood estimate, then an increasing value in the associated variable will lead to an increasing (decreasing) value of the forward intensity at that horizon, which in turn means an increasing (decreasing) value for the forward default probability at that horizon.

For the stock index one-year trailing return variable, most groups have default parameters that are slightly negative in the shorter horizons and then become positive in the longer horizons. When the equity market performs well, this is only a short-term positive for firms and in the longer term, firms are actually more likely to default. This seemingly counterintuitive result could be due to correlation between the market index and other firm-specific variables. For example, Duffie et al. [2009] suggested that a firm's DTD can overstate its creditworthiness after a strong bull market. If this is the case, then the stock index return serves as a correction to the DTD levels at these points in time.

As expected, we observe the different relationships between the short-term interest rate and default across economies. This observation possibly indicates different lead-lag relationships between credit conditions and the raising and cutting of short-term interest rates.

DTD is a measure of the volatility-adjusted leverage of a firm. Low or negative DTD indicates high leverage and high DTD indicates low leverage. Therefore, PD would be expected to increase with decreasing DTD. Indeed, the DTD level has negative default parameters across calibration groups.

Aggregate DTD can measure the overall degree of the volatility-adjusted leverage in an economy. As mentioned in Subsection 2.1, we use two kinds of sector–specific aggregate DTDs: one for financial firms, and the other for non–financial firms. In each economy, the default parameters for the two aggregate DTDs usually display different patterns. Such patterns may reflect different credit risk profiles of the economy–wide business environments.

The log ratio of the sum of cash and short-term investments to total assets (CASH/TA) measures liquidity of a financial firm. Likewise, the log ratio of current assets to current liabilities (CA/CL) stands for liquidity of a non–financial firm. These two ratios indicate the availability of a firm's funds and its ability to make interest and principal payments. On the whole, almost all economies have negative default parameters for such liquidity ratios, although the short–term and long–term effects differ across each calibration group.

The ratio of net income to total assets (NI/TA) measures profitability of a firm. The relationship between PD and NI/TA is as expected: the default parameters for NI/TA level is negative for all economies and all horizons.

The logarithm of the market capitalization of a firm over the median market capitalization of firms over the past one year within the economy (SIZE) does not have a consistent effect on PD across different economies. For example, in the US the default parameters for SIZE level are positive for almost all horizons, suggesting that the complexity of larger firms outweighs the potential benefits, such as diversified business lines and funding sources. On the other hand, in China the default parameters for SIZE level are negative across almost all horizons. The lack of similarity may reflect the different business environments in such respective economies.

The default parameters associated with DTD Trend, CASH/TA Trend, CA/CL Trend, SIZE Trend and NI/TA Trend are negative across almost all economies and horizons. The trend variables reflect momentum. The momentum effect is a short-term effect, and evidence of this is seen in the lower magnitude of the default parameters at longer horizons than at shorter horizons. The exception is the NI/TA Trend, which for some calibration groups has a higher magnitude at longer horizons.

The ratio of the individual firm's M/B to the economy M/B median (M/B) can either indicate the market mis-valuation effect or the future growth effect. This default parameter is negative for the US in the shorter term, indicating that higher M/B implies lower PD, and the future growth effect dominates during this period. On the other hand, in China and in the Developed Asia-Pacific calibration group, the default parameter for M/B is positive, indicating that for these economies, the market mis-valuation effect dominates.

Shumway [2001] argued that a high level of the idiosyncratic volatility (SIGMA) indicates highly variable stock returns relative to the market index, which is equivalent to highly variable cash flows. Empirically, the sign on SIGMA is different across countries and across prediction horizons.

4.2 Prediction Accuracy

In-sample testing: Various tests are carried out to test the prediction accuracy of the RMI-CRI PD forecasts. These tests are conducted in-sample.

A single calibration is conducted for the in-sample tests, using data until the end of the data sample. As an example, one-year PD forecasts are made for 31 December, 2000 by using the data at or before 31 December, 2000 and the parameters from the calibration. These PD forecasts can be compared to actual defaults that occurred at any time in 2001.

Accuracy ratio: The accuracy ratio (AR) is one of the most popular and meaningful measures of the discriminatory power of a rating system (BCBS, 2005). The AR and the equivalent Area Under the Receiver Operating Characteristic (AUROC) are described in Duan and Shrestha [2011]. In short, if defaulting firms had been assigned among the highest PD of all firms before they defaulted, then the model has discriminated well between safe and distressed firms. This leads to higher values of AR and AUROC. The range of possible AR values is in [0,1], where 0 indicates a completely random rating system and 1 stands for a perfect rating system. The range of possible AUROC values is in [0.5, 1]. AUROC and AR values are related by: AR = $2\times$AUROC - 1.

The AR and AUROC values for different horizons are available in Table B.1. Only economies with more than 20 defaults entering into the AR and AUROC computation are listed.

The AUROC values have been provided only for the purpose of comparison, if other rating systems report their results in terms of AUROC. The discussion will focus only on AR. The model is able to achieve strong AR results mostly greater than 0.80 at the one and six-month horizons for developed economies. There is a drop in AR at one and two-year horizons, but the AR are still mostly acceptable.

The AR in some emerging market economies such as China, India, Indonesia, and the Philippines are noticeably weaker than the results in the developed economies. This can be due to a number of issues. The quality of data is worse in emerging markets, in terms of availability and data errors. This may be due to lower reporting and auditing standards. Also, variable selection is likely to play a more important role in emerging markets. The variables are selected based on the predictive power in the US. Performing variable selections specific to the calibration group are expected to improve predictive accuracy, especially in emerging market economies. Finally, there could be structural differences in how defaults and bankruptcies occur in emerging market economies. If the judicial system is weak and there are no repercussions for default, firms may be less reluctant to default.

Aggregate defaults: The time series of aggregate predicted number of defaults and actual number of defaults in each calibration group are also available in Figs. B.3 to B.8. For India in particular, these figures show that there is room for improvement in the predictive power of the model.

5 Corporate Vulnerability Index

In July 2012, CRI launched the Corporate Vulnerability Index (CVI), which is a new suite of indices to produce bottom-up measures of credit risk in economies, regions and portfolios of special interest. The suite of CVIs is available in three distinctive types:

1. Value-weighted CVI (CVI_{vw}) RMI-CRI PDs are aggregated with each firm weighted by its market capitalization so that the size of each firm is taken into account.
2. Equally-weighted CVI (CVI_{ew}) RMI-CRI PDs are aggregated with each firm equally weighted. This captures the prevalence of credit risk by focusing on the number of firms at risk.
3. Tail CVI (CVI_{tail}) In taking the 5th percentile of the highest RMI-CRI PDs, the most vulnerable firms in a group are measured.

The CVIs are a set of indicators that gauge economic and financial environments in a new dimension. They are best viewed as stress indicators that reflect heightened credit risks in the corporate sector from three different angles.

Index Construction The primary inputs to the CVI are RMI-CRI 1-year PDs for individual exchange-listed firms.

- Value-weighted CVI (CVI_{vw}) CVI_{vw} is an aggregation of individual PDs weighted by each firm's market capitalization. In other words, at time t, given an interested group or portfolio G,

$$\text{CVI}_{\text{vw}}(t) = \sum_{i=1}^{I} \omega_{it} p_i(t, 12),$$

  where $p_i(t, 12)$ is firm i's default probability within 12 months viewed from t, $i \in \{1, 2, ..I\}$. Also, the weight for firm i at time t is ω_{it}, and $\omega_{it} = \frac{MC_{it}}{\sum_{i=1}^{I} MC_{it}}$, in which, MC_{it} is firm i's market capitalization at time t. If a firm does not trade on a particular day, the market capitalization from the previous valid day (within 20 trading days) is used. The market-capitalization weighting is applied to all economies and groups of economies, but is not applied to portfolios such as the S&P 500 index. The S&P 500 index is a float-adjusted index where the shares available to investors are used instead of the total shares outstanding, and our weighting scheme of CVI_{vw}(SPP) is consistent with the S&P 500 index.

- Equally-weighted CVI (CVI_{ew}) The equally-weighted CVI is computed by aggregating each firm's PD with equal weights applied to each firm. In other words,

$$\text{CVI}_{\text{ew}} = \frac{1}{I} \sum_{i=1}^{I} p_i(t, 12).$$

- Tail CVI (CVI_{tail}) The tail CVI provides a measure of the relatively more distressed firms in a group. It is the highest 5th percentile of PDs. The tail CVI can also be interpreted as the conditional median of the 10 percent tail, which is a more robust measure of "tail average" than the conditional mean of the 10 percent tail.

Inclusion of Firms: A firm's PD is computed with the model parameters from its primary exchange. The construction of CVI, however, is based on the firm's country of domicile. In regions like the Eurozone, some of the public holidays do not coincide. In this case, the aggregation is computed by using PDs from the previous trading day for firms that are listed in countries that have a public holiday, and PDs from the current trading day for firms that are listed in countries that do not have a public holiday. And firms are included in the Eurozone CVI only if their countries of domicile are part of the Eurozone at time t. For CVI of the S&P

500 portfolio, the constituents typically coincide with the constituents of the S&P 500 index for each point in time, and any missing PD value for a company in the S&P 500 is filled in with the most recently available PD.

6 Actuarial Spread

In July 2014, CRI launched a new credit risk measure, the Actuarial Spread (AS), which is the counterpart of market credit default swap (CDS) with contract horizons ranging from 1 year to 5 years but valued based on RMI-CRI's PDs in the forward horizons. Since then, the computation and publication of the AS have been implemented on a daily basis in addition to those of the PDs. Much like the par spread in a standard credit default swap (CDS) contract, the AS leverages the term structure of the physical PDs of the CRI and is essentially the premium rate that purely reflects the actuarial present value of a default protection. It provides a new metric of credit risk that the financial practitioners are more familiar with.

The construction of the AS relies on the features of a standard CDS contract. To fulfill a CDS contract, the protection buyer pays premiums on a regular basis to the seller until the contract matures or the reference entity defaults. In exchange, the protection buyer receives at the default time a contingent lump sum payment, the amount of which is based on the recovery rate of the reference instrument. Such a CDS contract terminates on its maturity date if there is no default up to its maturity; otherwise, it ceases on a default day, if any. Note that, if a default occurs during a payment period, the premium for the protection from the first accrual day to the default day is also assumed to be paid by the CDS buyer on the default day. Considering no effect from the market liquidity and using the physical PDs that CRI generates, the AS is calculated in a way that the expected present value of the contingent claim upon default is equal to the expected present value of the series of premiums up until the stop of a CDS contract. To familiarize the details of its theoretical formulation, please refer to Duan [2014]. As opposed to the continuous model introduced in Duan [2014], this technical report provides a discrete representation of the model for implementation purpose. For easy comparison, it adopts the same notations in the journal article as much as it possibly can.

A typical CDS contract adopts one day as the fundamental period of time. For this, we abbreviate the interval $((d-1)\cdot\Delta t,\, d\cdot\Delta t]$ in a forward time axis by the term day $d \in \mathbb{N}$ where $\Delta t = 1/365$ reflects the 365 day count convention. Consider t is the trading day of a CDS contract terminating on the day $T > t$. If the reference entity defaults at a random day τ where $t+1 \leq \tau \leq T$, he will in return get a lump sum payment, which is 1 minus the recovery rate R_τ, from a unit-notional CDS and cease to make the scheduled payment beyond the default point. We assume the premiums are scheduled to be paid on the days $t_1, t_2, ..., t_k$ with $t_k = T$, where each payment period is roughly three months. Note that a payment day t_{i-1} is also the first day of the coming accrual period, which ends on the day before next payment day, denoted and defined by $t'_i = t_i - 1$. However, a trading day t may also occur after a payment day, say t_{i-1}, and we denote the exact start date of its remaining accrual period by $t_{i-1} \vee (t+1) = \max\{t_{i-1}, t+1\}$ for a general purpose.

Another actual/360 day count convention is usually adopted to define the length in year of an accrual period, for which we denote $A(s,q)$ the period length in year from the day s to the day $q > s$ (both inclusive). For example, if a quarterly accrual period from t_{i-1} to t'_i (both inclusive) has 91 days, then $A(t_{i-1}, t'_i) = 91/360$ is applicable.

Compared to the risk-neutral probability measure used in the CDS pricing, the AS is essentially its counterpart based on a physical probability measure P. We denote it by $S_t^{(a)}(T-t)$ with its days to maturity $(T-t)$. Following the assumption that there is no arbitrage for CDS buyer and seller, the AS is defined to satisfy the equation:

$$
\begin{aligned}
& E_t^p\left[(1-R_\tau)D_t(\tau-t)\cdot\mathbb{1}_{\{t<\tau\le t'_k\}}\right] \\
= \quad & S_t^{(a)}(T-t)\sum_{i=1}^{k}\Big\{ \qquad A(t_{i-1}\vee(t+1),t'_i)\cdot E_t^p\left[D_t(t_i-t)\cdot\mathbb{1}_{\{t'_i<\tau\}}\right] \\
& \qquad\qquad +E_t^p\left[A(t_{i-1}\vee(t+1),\tau)\cdot D_t(\tau-t)\cdot\mathbb{1}_{\{t'_{i-1}<\tau\le t'_i\}}\right]\Big\},
\end{aligned}
$$

where E_t^P is an expectation operator with respect to the physical probability measure P, τ refers to the random default day, $D_t(\tau-t)$ is the random money market discount factor starting from the day t to another day τ and k is the number of the CDS premium payments.

The real-time LIBOR rates up to one year and Swap rates beyond are generally available from the market. With the combination, one can bootstrap the implied LIBOR rates beyond one year. As the AS is calculated based on days, a linear interpolation is further performed to obtain the implied LIBOR rates up to each forward day (in continuously compounded annualized form), which then serve the role of the discount factor $D_t(\cdot)$. Let $r_t(s,q)$ be the day-t risk-free annualized forward discount rate between the day $t+s$ and the day $t+q$ (both inclusive) with $q\ge s\ge 1$. In particular, $r_t(1,q)$ refers to the day-t risk-free spot discount rate covering the days $t+1,\ldots,t+q$. The standard term structure theory implies that

$$
r_t(1,q)=-\frac{1}{q}\ln\left(E_t^P\left[D_t(q)\right]\right).
$$

Further we let $r_t(q,q)=r_t(1,q)\cdot q-r_t(1,q-1)\cdot(q-1)$ for $q\ge 2$, which refers to the day-t instantaneous forward rate for the day $t+q$. As will be seen later, defining $r_t(s,q)$ this way is to make it consistent with the definition of the forward default/other exit intensity in terms of the day count convention. With the RMI-CRI PDs serving as the physical probability measure P and the use of a standard recovery rate of $\bar{R}_t=40\%$, the AS is rewritten as

$$
S_t^{(a)}(T-t)=
$$

$$
\frac{(1-\bar{R}_t)\cdot E_t^p\left[e^{-r_t(1,\tau-t)(\tau-t)/365}\cdot\mathbb{1}_{\{t<\tau\le t'_k\}}\right]}{\sum_{i=1}^{k}\left\{A(t_{i-1}\vee(t+1),t'_i)\cdot e^{-r_t(1,t_i-t)(t_i-t)/365}\cdot E_t^p\left[\mathbb{1}_{\{t'_i<\tau\}}\right]\;+\;E_t^p\left[A(t_{i-1}\vee(t+1),\tau)\right]\cdot e^{-r_t(1,\tau-t)(\tau-t)/365}\cdot\mathbb{1}_{\{t'_{i-1}<\tau\le t'_i\}}\right\}}, \quad (37)
$$

where the actual/365 day count convention is used for the discount factor and integration.

To obtain the physical probability of defaults and their term structures, we apply CRI's forward intensity model. Define $f_t(u)$ to be the day-t forward default intensity over the day $t+u$, which will be used to calculate the probability of default of a firm conditioning on its survival up to the day $t+(u-1)$. The forward intensity for other exits, or $h_t(u)$, can be similarly defined. These two intensities are expressed as exponential linear functions of 17 variables in general, including an intercept term, 4 common covariates and 12 firm-specific covariates, in the form of

$$
f_t(u)=\exp\{\alpha_0(u)+\alpha_1(u)x_{1,t}+\ldots+\alpha_{16}(u)x_{16,t}\},
$$

and

$$
h_t(u)=\exp\{\beta_0(u)+\beta_1(u)x_{1,t}+\ldots+\beta_{16}(u)x_{16,t}\}.
$$

In this similar manner, 15 variables for China apply to the two intensities (see Subsection 2.1). The coefficients $\alpha_i(u)$ and $\beta_i(u)$ are functions of forward starting time, which are further modelled by Nelson-Siegel term structure functions, such as

$$
\alpha_i(u;\varrho_{i,0},\varrho_{i,1},\varrho_{i,2},d_i)=\varrho_{i,0}+\varrho_{i,1}\frac{1-\exp(-u\Delta t/d_i)}{u\Delta t/d_i}+\varrho_{i,2}\left[\frac{1-\exp(-u\Delta t/d_i)}{u\Delta t/d_i}-\exp(-u\Delta t/d_i)\right], \quad (38)
$$

for $i = 0, 1, 2, \ldots, 16$. Recall that, except for the intercept terms $\alpha_0(u)$ and $\beta_0(u)$, the other covariates are stochastic and their long-term levels are restricted to zeros; namely, $\varrho_{i,0} = 0$ for $i = 1, 2, \ldots, 16$. With $f_t(u)$ and $h_t(u)$ in place, we are ready to define $\psi_t(s,q) = \frac{\sum_{u=s}^{q}[f_t(u)+h_t(u)]}{q-(s-1)}$, for $q \geq s \geq 1$, which is a standardized forward termination intensity covering the days $t+s, \ldots, t+q$.

One important feature of the CDS is that when the reference entity ceases to exist due to reasons other than default, such as mergers and acquisitions, the CDS protection is typically shifted to the merged or acquiring entity. Naturally, we should take into account the fact that the successor entity will then face subsequent default or other exits. There indeed are a number of ways to model the relationship between the termination probability of the reference entity and the successor entity (see [Duan, 2014]). In CRI's implementation, we further assume that the successor has the forward default and other exit intensities identical to those of the original reference entity.

Let $P_t(s, q; r_t(1,u), s \leq u \leq q)$ denote the day-t discounted forward probability of the reference entity of the CDS being terminated, including successions, over the days $t+s, \ldots, t+q$. Under the assumptions above, Duan [2014] has derived its analytical solution, which can be re-written in the discrete form below

$$P_t(s, q; r_t(1,v), s \leq v \leq q) = \sum_{v=s}^{q} e^{-\sum_{u=s}^{v}[r_t(u,u)+f_t(u)]\Delta t} f_t(v)\Delta t. \tag{39}$$

By temporarily setting the forward interest rate to 0 in Eq. (39), the first term of denominator in Eq. (37) can be presented in the form of

$$E_t^P(1_{\{t_i' < \tau\}}) = 1 - P_t(1, t_i' - t; r_t(1,u) = 0 \text{ for } 1 \leq u \leq t_i' - t). \tag{40}$$

The solutions to the two remaining two terms of Eq. (37) can be expressed as

$$\begin{aligned}
& E_t^p\left[e^{-r_t(1,\tau-t)(\tau-t)/365} \cdot \mathbb{1}_{\left\{t<\tau\leq t_k'\right\}}\right] \\
= \quad & \sum_{q=1}^{t_k'-t} e^{-[r_t(1,q)+\psi_t(1,q)]\cdot(q/365)} \cdot f_t(q) \cdot \Delta t \\
+ \quad & \sum_{q=1}^{t_k'-t} e^{-[r_t(1,q)+\psi_t(1,q)]\cdot(q/365)} \cdot h_t(q) \cdot P_t(q, t_k' - t; r_t(1,v), q \leq v \leq t_k' - t) \cdot \Delta t
\end{aligned}$$

and

$$
\begin{aligned}
& E_t^p\left[A(t_{i-1}\vee(t+1),\tau)\right]\cdot e^{-r_t(1,\tau-t)(\tau-t)/365}\cdot \mathbb{1}_{\{t'_{i-1}<\tau\leq t'_i\}} \\
= & \sum_{q=t_{i-1}\vee(t+1)}^{t'_i} A(t_{i-1}\vee(t+1),q)\cdot e^{-[r_t(1,q-t)+\psi_t(1,q-t)]\cdot(q-t)/365}\cdot f_t(q-t)\cdot\Delta t \\
+ & \sum_{q=t_{i-1}\vee(t+1)}^{t'_i} A(t_{i-1}\vee(t+1),q)\cdot e^{-[r_t(1,q-t)+\psi_t(1,q-t)]\cdot(q-t)/365}\cdot h_t(q-t) \\
& \cdot P_t(q-t,t'_i-t;r_t(1,v),q-t\leq v\leq t'_i-t)\cdot\Delta t
\end{aligned}
$$

With the formulas mentioned above, we compute the AS, or $S_t^{(a)}(T-t)$, and provide it to the public on a daily basis.

7 CriSIFI

In August 2017, CRI launched the CRI Systemically Important Financial Institution (CriSIFI) on its website (`http://rmicri.org`). Further, we updated the system twice in September 2017 and January 2018, which enables users to assess systemic importance of exchange–listed banks and insurers globally. The CriSIFI aims to identify systemic risks of those banks and insurers by capturing their tendency to default together (i.e., too connected to fail) along with their respective asset sizes (i.e., too big to fail). For example, a financial institution with a higher ranking (e.g., 10 is a higher ranking than 20) is likely to pose a higher risk to the financial system and thus has greater systemic importance than does a lower ranked firm. In short, the CriSIFI relies on a novel way to construct a proper financial network which combines nodes and edges of a network.

- Node: firm characteristics captured by the ratio of individual financial institution's assets over the network's total assets
- Edge: network configuration reflected through partial default correlations of financial institutions

The CriSIFI data panel is monthly updated and starts from January 2000. The CriSIFI is updated monthly on the CRI website where all exchanged-traded banks (banks and investment banks) and insurers globally are included. For details, see Table A.1 for the CRI coverage. The CriSIFI can be used to track and monitor systemic risk of each financial institution in the global financial system. Apart from the CriSIFI, the CRI reports the CRI Systemically Important Bank (CriSIB) and the CRI Systemically Important Insurer (CriSII) globally, or within a local community such as region (e.g., North America and Asia-Pacific Developed economies) and economy (e.g., U.S. and Singapore). All three systemic importance indicators can help identify potential systemic risk via financial institutions' connectedness in the global financial network. Next, we explain how to construct the CriSIFI.

7.1 Constructing the forward-looking PD partial correlation matrix

A primary input to the CriSIFI is the forward–looking PD (probability of default) partial correlation matrix, which is used to measure connectedness between financial institutions in the network. This partial correlation matrix is generated from the forward–looking PD total correlation matrix using the model of Duan and Miao [2016], which is a factor model along with sparsely correlated residuals for PDs and POEs (probabilities of other exist) of all firms considered. It is worth noting that POE is a crucial element for properly estimating multiple–period default probabilities, because suitable survival probability of a firm in a multiperiod context cannot be determined without POE (see Duan et al. [2012]). Omitting POE is particularly troublesome when knowing that POEs are empirically many folds larger than PDs. First, we briefly explain how to obtain the forward–looking PD total correlation matrix. It is important to note that our methodology follows that of Chan-Lau et al. [2016], which is largely based on Duan and Miao [2016] except for deploying a logit transformation instead of a double–log transformation.

(a) Define one pair of predetermined global factors, ten pairs of predetermined industry factors, and one pair of predetermined economy factors for each economy of domicile (one–month, logit–transformed, median PD and POE). The logit transformation, denoted by a hat, has the following form:

$$\widehat{PD} = \log \frac{PD}{1 - PD} \text{ and } \widehat{POE} = \log \frac{POE}{1 - POE}.$$

The logit transformation is valid because PDs and POEs all fall in (0,1). A dynamic model is then constructed on these 24 $\widehat{PD}$ and $\widehat{POE}$ factors. Later, the inverse transformation will be applied to recover simulated model PD and/or POE factors:

$$PD = \frac{\exp(\widehat{PD})}{1 + \exp(\widehat{PD})} \text{ and } POE = \frac{\exp(\widehat{POE})}{1 + \exp(\widehat{POE})}.$$

(b) In particular, the predetermined economy pair should have at least 30 observations available in the domicile economy. Otherwise, we use the median PD/POE pair of aggregation groups as a substitution: Asia Pacific (Developed), Asia Pacific (Emerging), Europe, Latin America & Caribbean, Sub–Saharan Africa, or Middle East, North Africa & Central Asia. In case an economy has sufficient observations (equal or more than 30) in the history but not later on, we continue to use the economy median. If the economy has fewer observations earlier but sufficiently large later on, we allow the switch from the group median to the economy median to happen but for only once.

(c) The global pair of $\widehat{PD}$ and $\widehat{POE}$ are normalized to have mean 0 and variance 1. For each industry factor, regress $\widehat{PD}$ (or $\widehat{POE}$ factor) on the pair of the global factors to remove any shared information arising from the global factors (i.e., orthogonalization). Henceforth, the industry factors refer to the "orthogonalized regression residuals" uncorrelated with the global factors. We then normalize the 10 industry pairs of $\widehat{PD}$ and $\widehat{POE}$ residuals and the 1 predetermined pair of $\widehat{PD}$ and $\widehat{POE}$ to have a standard deviation of 1 (i.e., normalization).

(d) Model the factors with a bivariate vector autoregressive process of order one without intercept terms, i.e., VAR(1), for each of the 12 pairs of $\widehat{PD}$ and $\widehat{POE}$ factors by deploying entire historical data series up to the point of analysis. Doing so ensures that the factor dynamics are estimated with data covering different phases of a credit cycle and over several credit cycles. Note that the intercept terms are set to zero because normalization has removed the mean.

(e) Estimate the "best" factor model by regressing individual firm $\widehat{PD}$ on 12 global, industry, and economy $\widehat{PD}$ factors using a 60–month moving data window. Likewise, regress individual firm $\widehat{POE}$ on 12 global, industrial, economy $\widehat{POE}$ factors. Deploy the adaptive lasso technique of Zou [2006] with cross–validation in these regressions to avoid overfitting.

(f) Individual firm's factor model residuals (60 data points at most) are treated as an AR(1) process, and the AR residuals are then used to compute cross–firm correlations. Note that some individual firm's $\widehat{PD}$ and $\widehat{POE}$ are missing due to bankruptcies and/or mergers/acquisitions. We thus construct the AR residual correlation matrix by first computing pairwise correlations, and then apply thresholding coupled with cross–validation to identify a legitimate "sparse" AR residual total correlation matrix.

(g) Use the estimated factor model along with sparse residual correlations to simulate future PDs and POEs for all financial institutions under consideration, and with which we can apply the survival/default formula on the simulated PDs and POEs to obtain PD over any prediction horizon of interest via Monte Carlo averaging of the stochastic PD term structure for each financial institution. This theoretical PD term structure under a particular parameter value serves as the basis to recalibrate factor loadings for every financial institution via a single firm–specific scaling factor and the parameters of its residual AR(1) model. Our recalibration is implemented to fit the 5–year PD term structure provided by the CRI system. This recalibration step ensures that default correlations are obtained not at the expense of poorly matching the available PD term structure individually.

(h) Use the recalibrated model to simulate PDs and POEs for a specific horizon of interest (e.g., one year) at any future time point (e.g., one month later), and estimate the forward–looking total default correlation matrix using the simulated sample.

Importantly, we focus on the forward–looking default correlation via simulation, not on the historical average available from the time series of PDs in the CRI database. The reason is that this average measure represents backward-looking comovements, which does not represent the future when one goes through different phases of a credit cycle. In contrast, the forward-looking correlations reflect the currently available information and should better gauge the potential riskiness going forward. Readers who are interested in comparing the forward-looking and backward-looking results are referred to Chan-Lau et al. [2016]. Other practical considerations also favor forward-looking default correlations over historical default correlations. For example, considering 1–year PD correlations over a period of six months instead of one month would see a dramatic reduction in usable sample size by a factor of six.

Apart from the use of the forward–looking PDs, we focus on "partial" not "total" correlations. Partial correlation is the residual correlation after removing any indirect connections through other parties in the network. Conceptually, partial correlation rightfully captures the direct default connection between any two financial institutions. Of course, indirect connections are also of interest for network analysis, but they are already reflected through the network configuration represented by many direct bilateral linkages. We obtain the partial default correlation matrix through a regularization technique.

We use the CONCORD (CONvex CORrelation selection methoD) algorithm of Khare et al. [2015] and Oh et al. [2014]. Conceptually, it amounts to imposing zero partial correlations on pairs with weak ties. The CONCORD algorithm also ensures convergence because it preserves convexity through an appropriate selection of weights and a particular design of the penalty term on the concentration matrix rather than on the partial correlation matrix. In addition, the high dimensional data calls for regularization, simply because high dimensionality left un-regularized may deliver a highly unstable partial correlation matrix. As a result, the globally connected and regularized network will be more stable and does not generate an overwhelmingly large number of systemic firms.

Specifically, the CONCORD objective is to minimize

$$Q_{con}(\Omega) = \frac{N}{2}\left[-\ln\left[\det(\Omega_D^2)\right] + \mathrm{tr}(S_N\Omega^2) + \lambda||\Omega_X||_1\right],$$

where $\det(\cdot)$ denotes the determinant operator; $\mathrm{tr}(\cdot)$ denotes the trace operator; S_N is the sample correlation matrix computed with a sample size of N; $\Omega = \Omega_D + \Omega_X$ is the concentration matrix (i.e., the inverse of the correlation matrix); $\lambda > 0$ is the tuning parameter used to determine the shrinkage rate or how aggressively one penalizes the non–zero entries in Ω_X; $\lambda||\Omega_X||_1 = \lambda \sum_{i \neq j} |\omega_{ij}|$ is the L_1–penalty term; and ω_{ij} is the off–diagonal element in Ω_X. Here, we select a λ below which totally isolated firms in the network begin to emerge. The tolerance error for finding the optimal λ and the partial correlation precision are respectively set to 10^{-3} and 10^{-4}. For technical details, see Chan-Lau et al. [2016].

7.2 Computing the CriSIFI

The CriSIFI is a network centrality indicator used to assess the relative importance of a financial institution in the network, and is the appropriate entry in the non-negative eigenvector of $Q|\bar{P}_{X,t}|Q$ that corresponds to the largest eigenvalue. $|\bar{P}_{X,t}|$ is the absolute value of $\bar{P}_{X,t}$ and $\bar{P}_{X,t}$ denotes the 12–month moving average of $P_{X,t}$, the regularized partial correlation matrix at time t after setting its diagonal elements to 0. Deploying the 12-month moving average is to remove the excessive noise. Q is a diagonal matrix with q_i as its i–th diagonal element where q_i is the size of a financial institution over the total size of the network, measured in USD; Technically, $Q|\bar{P}_{X,t}|Q$ is a non–negative matrix, and the Perron–Frobenius theorem ensures the existence of such a non–negative eigenvector.

The CriSIFI captures both the node (the firm's asset size) and edge (the strength of connectedness reflected in the partial correlation) characteristics in the financial network. We contend that our forward-looking systematic risk ranking, combining both the edge and node characteristics, is much more comprehensive than the alternatives: (1) a backward-looking ranking measure, and (2) any measure that only factors in one of the two characteristics. Therefore, under the CriSIFI small financial institutions being connected to large ones may present significant systemic risks simply due to the feedback effect from their connected larger counter–parties. Chan-Lau et al. [2016] also compare the performance of the CriSIFI with those of other measures such as Global Systemically Important Banks (G-SIBs) released by the Financial Stability Board (FSB). They find that the G-SIBs are likely to be biased toward singling out large financial institutions in the system, and overall connectivity only plays a rather minor role.

8 Ongoing Developments

The CRI can develop a number of directions. We now comment on obvious ones that in our view are likely to bring meaningful and measurable benefits. Besides modifications to the current modeling framework of the forward intensity, a change in modeling platform will be undertaken if another model proves more promising in terms of accuracy and robustness of results. For this type of development, we also rely on the collective efforts by the worldwide credit research community to challenge and improve the existing modeling platform.

Within the current modeling framework, future developments involve, for example, the CRI plans to implement DTD estimations by a novel density-tempered expanding-data sequential Monte Carlo method. Another challenging example includes variable and structural–break selections where Artificial Intelligence automatically identifies time window, crucial risk factors, and breakpoints regarding defaults in a way that we would consider "smart". Also, we are designing a more comprehensive treatment scheme to handle missing data.

Finally, a series of new applications and tools using the RMI-CRI PDs as an input are currently being developed. More specifically, the CRI is actively working with users and exploring different possibilities of taking advantage of the world-class research infrastructure at the institute to propagate real world applications in credit rating and testing. The CRI has developed a tool for stress testing the financial stability for economies around the world. The CRI has also developed a methodology to address default correlations within a portfolio. The CRI remains committed to making its vast resources available for academic research.

Acknowledgements

The RMI Credit Research Initiative is premised on the concept of credit ratings as a "public good". Being a non-profit undertaking allows a high level of transparency and collaboration that other commercial credit rating systems cannot replicate. The research and support infrastructure is in place and researchers from around the world are invited to contribute to this initiative. Any methodological improvements that researchers develop will be incorporated into the CRI system. In essence, the initiative operates as a "selective wikipedia" where many can contribute but implementation control is retained.

If you have feedback on this technical report or wish to work with us in this endeavor, please contact us at rmicri@nus.edu.sg.