Publication Date: 16 Oct 2019

Effective Date: 14 Oct 2019

Addendum 13 to the CRI Technical Report (Version: 2017, Update 1)

This document updates the Technical Report (Version: 2017, Update 1) and details the method handling some partially observed default events in India.

TransUnion Credit Information Bureau India Limited (CIBIL)[1] has publicly released the list of Indian defaulted firms in each quarter from Q1/2001 onward. This additional data source enriches the CRI default database, and its net effect is to significantly increase the number of defaults from original 512 to 1,336. However, the CIBIL default database only provides the calendar quarter of a default. In order to utilize these default events, the CRI team has to modify the likelihood function to estimate the default intensity parameters. Per the effective date stated in this document, the CRI-PD model for India has reflected the change in its estimation methodology that incorporates these CIBIL reported default events. The original and revised results are displayed in Figure 1.

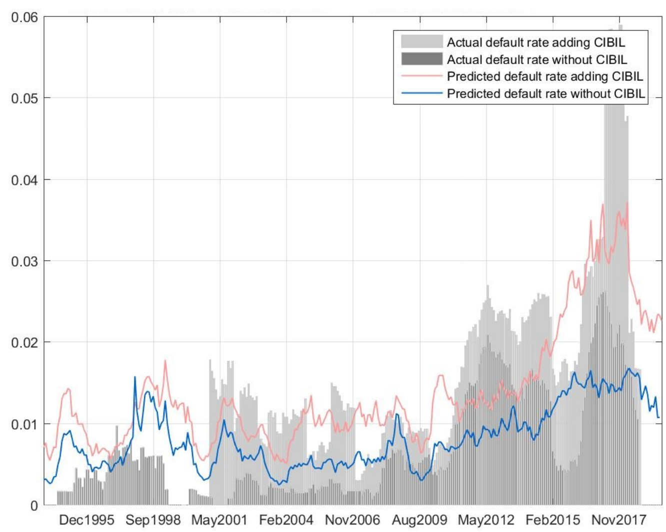

Figure 1: Comparison between actual and predicted 12-month default rates

[1]https://www.cibil.com

I.Likelihood function for Indian default data

The CRI-PD model employs a likelihood function based on the forward-intensity approach of Duan, Wang, and Sun (2012)[2], which basically comprises two independent components governing separately default and other corporate exits. The model is estimated to the data on a monthly frequency. Since the default events reported by CIBIL are calendar quarters of their occurrences, the likelihood function must be modified to reflect these partially observed information.

The data is partitioned to reflect two types of default information:

$$\mathbf{X}_N = (X_n, X_m),$$

where $X_n = \{x_i\}_{i=1}^{n}$ are the variables corresponding to an i-th firm whose default date (τ_{D_i}), combined exit date[3] (τ_{C_i}), or survival is fully observed, whereas $X_m = \{x_j\}_{j=1}^{m}$ represent the variables of defaulted firms reported in the CIBIL's default data and yet not in the CRI's original default list. In other words, we only know their default quarter, i.e., $\tilde{\tau}_{D_j}$ is partially observed, and likewise their combined exits quarter $\tilde{\tau}_{C_j}$, $i \neq j$. Naturally, $n + m = N$ is the total number of the Indian firms.

The likelihood function for the default intensity parameters' estimation can be written as:

$$\mathscr{L}_{\tau}^{*}(\alpha; \tau_C, \tau_D, \tilde{\tau}_D.\tilde{\tau}_C, \mathbf{X}_N) = \prod_{t=0}^{T-1} \Big[\prod_{i=1}^{n} \mathscr{L}_{\tau,i,t}(\alpha; \tau_{D_i}, \tau_{C_i}, x_{it}) \prod_{j=1}^{m} \tilde{\mathscr{L}}_{\tau,j,t}(\alpha; \tilde{\tau}_{D_j}, \tilde{\tau}_{C_j}, x_{jt}) \Big], \tag{1}$$

where τ is the maximum prediction horizon[4], α is the set of unknown parameters for the default intensity functions, $\mathscr{L}_{\tau,i,t}(\cdot)$ is the original decomposed likelihood function[5] applied to all i's, and $\tilde{\mathscr{L}}_{\tau,j,t}(\cdot)$ is a modified term to reflect partially observed default events for all j's.

Let t_j^* be the last month of the quarter preceding the default quarter. The modified likelihood function below reflects the fact that at $t_j^* + 3$ the exact default month is still unknown, but can be assessed by, say, applying a prior probability distribution over the three months in a defaulting quarter, tallied from other default cases for which we know their defaulting months. Denote this prior probability distribution by w_1, w_2 and w_3.

There are two cases to deal with for this subset of CIBIL reported default firms. We need to set $\tau_{D_j} = \tau_{C_j} = t_j^* + 3$ to reflect the fact that the default event becomes known only at the end of

[2]Duan, J. C., Sun, J., & Wang, T. (2012). Multiperiod corporate default prediction – a forward intensity approach. *Journal of Econometrics*, 170(1), 191-209.

[3]Combined exits including default exit and other corporate exits

[4]The maximum prediction horizon is 60 months in the CRI implementation

[5]See equation (11) in Duan, Wang, and Sun (2012) for further detail

a default quarter. Thus,

$$\begin{aligned}\tilde{\mathscr{L}}_{\tau,j,t}(\alpha;\tilde{\tau}_{D_j},\tilde{\tau}_{C_j},x_{jt}) &= 1_{\{t_{0j}\le t\le t_j^*+3\ \&\ t+\tau<t_j^*+3\}}\mathscr{L}_{\tau,j,t}(\alpha;\tau_{D_j},\tau_{C_j},x_{jt}) \\ &\quad +1_{\{t_{0j}\le t\le t_j^*+3\ \&\ t+\tau=t_j^*+3\}}\sum_{k=1}^{3} w_k P_t^D\{\tilde{\tau}_{D_j}=t_j^*+k\}+1_{\{t_{0j}>t\}}, \qquad (2)\end{aligned}$$

where $P_t^D\{\tilde{\tau}_{D_j} = t_j^* + k\}$ denotes the probability at the prediction time t for firm j defaulting at $t_j^* + k$, which can be computed per the forward intensity model. The second term on the right-hand side reflects the fact that we do not know the exact default month after a default has been reported. The best one can do is to weight the three default probabilities with the three prior month distribution probabilities. The third item in the equation basically addresses firms for which data only become available after the prediction time.

In terms of the CRI implementation, we specify the weight w_k in (2) as the proportion of the observed default months in a quarter. For example, as observed in September 2019, 46% of the total default firms is in the third month of the defaulting quarter, whereas each of the remaining two months constitutes approximately 27%. For the CRI website displays, an exact default date for the CIBIL firm is placed as the average day based on the observed sample distribution over days in a defaulting quarter.