Publication Date: 23 February 2017
Effective Date: 23 February 2017
Addendum 1* to the CRI Technical Report (Version: 2016, Update 1)

This document updates the Addendum 1 of the CRI Technical Report (Version: 2016, Update 1) and details a further improvement to the structural break specification of the CRI PD model for the Chinese sample vis-a-vis the one in Addendum 1. This change has been implemented for the Probabilities of Default (PDs) and Actuarial Spreads (ASes) released on 23 February 2017. This change has simplified the model specification while delivering a better default prediction performance for the Chinese firms. The way to compute the statistical inference is also explained in this updated addendum.

I. Revision to the parameter specification on the intercept and the DTD Level

In the CRI PD model, the parameters for all covariates are assumed to be time-invariant. The overall performance of the model on various countries/groups (e.g. North America, Europe, etc.) is good (e.g. the accuracy ratio (AR) for 1-year PD for the North American firms is 82%; the 1-year AR for emerging market is 77%). Before introducing the structural break specification in Addendum 1, one rare exception was the Chinese sample with an AR of 62% for 1-year PD. The relatively poor performance was also evident in Figure 1 using the Chinese sample up to 31 January 2017, which shows the predicted number of defaults over the 1-year horizon under the original model has missed the general pattern of the realized number of defaults over the prediction horizon for the sample period.

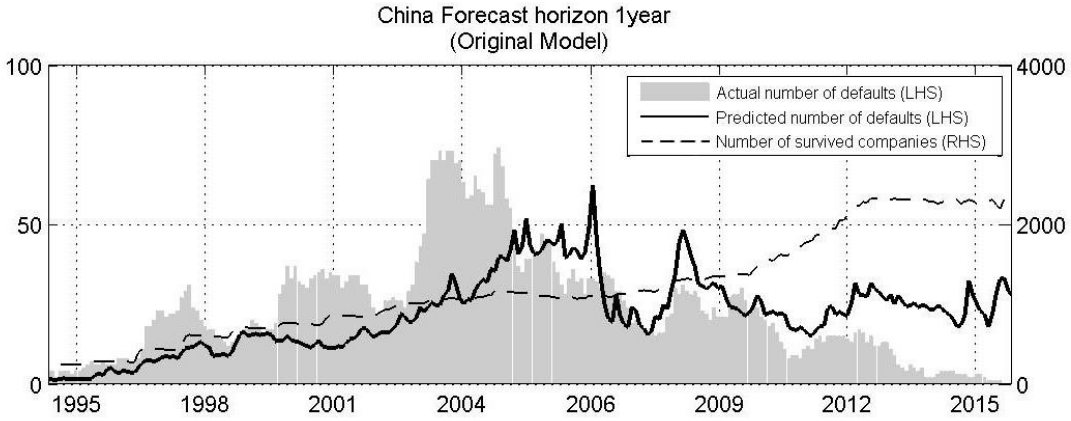

Figure 1

With a structural break in December 2004 being added to two parameters (i.e., coefficients for the intercept and DTD Level), the CRI model's performance on Chinese firms has been measurably improved. The adopted modification is to allow the break to occur in a smooth fashion instead of using a 0 and 1 dummy, and the smooth transition is accomplished by employing a logistic function.

Denote by t_0 the default prediction time where a structural break occurred, which is set to be December 31, 2004 for Chinese firms. We define the parameter for prediction horizon τ which is subject to a structural break at t_0 as:

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

where $\tilde{\alpha}(\tau)$ is a positive function of τ, controlling the rate of transition from one parameter value to another. For each prediction horizon τ, $\beta(t,\tau;t_0)$ moves in a smooth manner from $\tilde{\beta}(\tau)$ to $\tilde{\beta}(\tau) + \tilde{\gamma}(\tau)$ when the default prediction time advances toward and then beyond t_0. The specification of $\tilde{\beta}(\tau)$ follows the original model where the four-parameter and three-parameter Nelson-Siegel (NS) functions are used for the intercept term and the coefficient of DTD level, respectively (CRI Technical Report, Version: 2016, Update 1). All functions of τ in the second term on the right-hand side of the above equation are modeled by a three-parameter NS function. Using three parameters ensures that the NS function approaches zero as τ goes to infinity.

The model specification in Addendum 1 has 52 parameters in total where 12 are the newly introduced and 40 are in the original model for Chinese firms. The model with a structural break in two parameters delivers a far superior performance on predicting defaults of Chinese firms. The accuracy ratios for 1-year, 2-year, and 5-year PDs have increased from 59%, 52% and 39% to 67%, 64%, and 52%, respectively. Figure 2 using the updated Chinese sample, i.e., up to 31 January 2017, documents a marked improvement of the structural break model in terms of predicting the number of defaults over the original specification with its performance plotted earlier in Figure 1.

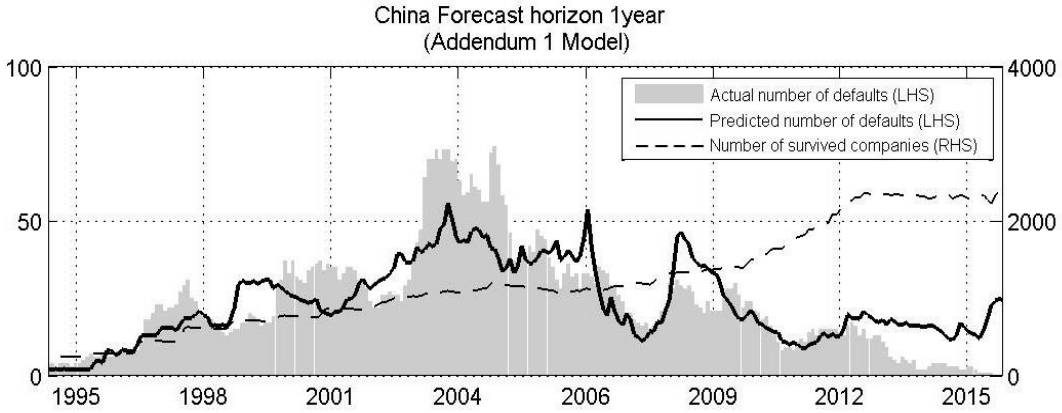

Figure 2

In this revised addendum, a further modification to the model specification is introduced. First note that the NS function contains four parameters: ρ_0, ρ_1, ρ_2 and *d*, which have been previously defined in the CRI Technical Report (Version: 2016, Update 1). For the CRI-PD model's intercept term, ρ_1 in the NS functions for $\tilde{\beta}(\tau)$ and $\tilde{\gamma}(\tau)$ is set to 0. However, ρ_0 in the NS functions for $\tilde{\beta}(\tau)$ and $\tilde{\gamma}(\tau)$ is kept free so as to allow for a permanent effect, i.e., the default prediction horizon approaches infinity. For the coefficient on the DTD level, the revised model continues to deploy a three-parameter NS function where ρ_0 is set to 0. The revised model has 51 parameters in total where 12 are added due to the structural break and the 40 parameters in the original model are reduced to 39. These parameters are again estimated with the sequential Monte Carlo method as described in the CRI Technical Report (Version: 2016, Update 1).

The new model specification with 51 parameters has delivered a similar goodness of fit by comparing Figures 2 and 3, which are generated with the Chinese sample up to 31 January 2017. The new model has comparable accuracy ratios of 69%, 66%, and 54% for 1-year, 2-year, and 5-year prediction horizons, respectively, as compared to the structural break specification prior to the revision. However, the new specification delivers notable improvement for shorter prediction horizons; for example, the accuracy ratio for the 1-month PD has increased from 52% to 70%.

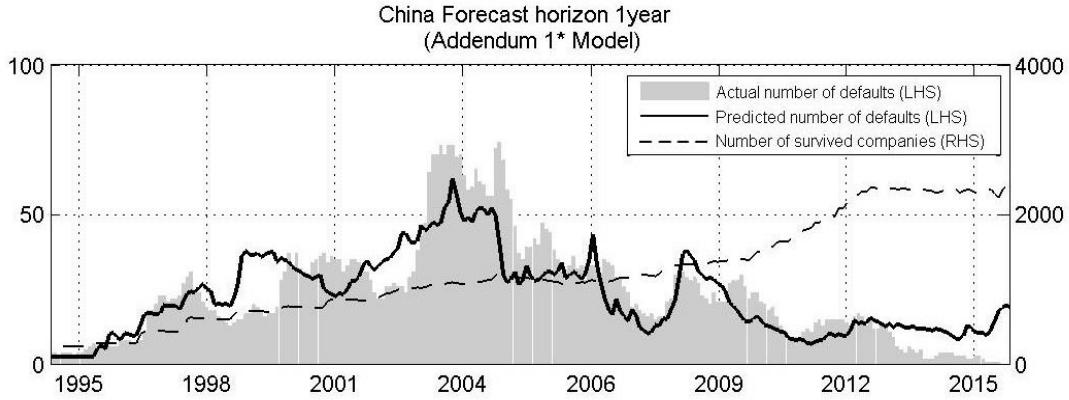

Figure 3

II. Statistical inference of the parameters with structural break

The statistical inference on the parameters for the intercept and DTD level are again based on Shao's self-normalized statistic (see CRI Technical Report, Version: 2016, Update 1 for details). Since the parameters in connection with the structural break cannot be identified using the data prior to the break point, the sequence of parameter estimates used in Shao's self-normalized statistic can only start from the break point onward. In our implementation, all parameter estimates, break or non-break related, start from the break point. Denote by T the endpoint of the dataset 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 (CRI Technical Report 2016, equation (28)) therefore equals $T - t_0 + 1$.