Publication Date: December 18, 2017 (original) and December 22, 2017 (revised)
Effective Date: December 18, 2017

Addendum 5 to the CRI Technical Report (Version: 2017, Update 1)

This addendum presents two changes to the pseudo-Bayesian parameter estimation method used in the CRI system previously documented in the Technical Report Version: 2017, Update 1:

- Remove the prior from the pseudo-Bayesian optimization algorithm; that is, using a uniform or improper prior (i.e., a constant) to replace the previous normal/truncated normal prior.
- Relax some restrictions on the Nelson-Siegel (NS) parameters concerning the structural-break model for China.

I. Remove the prior

Starting from Technical Addendum 4 of 2017, the CRI system has contained at least 46 NS parameters corresponding to one intercept term and 14 covariates. The number of NS parameters can be even larger when applying a smoothly-transitioned structural break for an economy, and with which the forward-intensity model's parameters will vary over time. Obviously, the estimation of the many parameters presents a numerical challenge.

To confront this challenge, we apply the density-tempered Sequential Monte Carlo (SMC) method of Duan and Fulop (2013) which is based on a pseudo-Bayesian principle. The pseudo-Bayesian estimation relies on the following typical Bayesian setup:

$$\text{Pseudo-Posterior} \propto \text{Pseudo-Likelihood} \times \text{Prior}.$$

The knowledge of the prior, which is taken as one's best guess, helps find the true (posterior) parameters. In the previous CRI implementation, for example, the prior (precisely, normal/truncated normal prior) of the NS parameters was obtained from fitting β (PD model parameters from the Maximum-Likelihood-Estimation (MLE) approach) into the specified NS function. In the revised CRI implementation, the prior has been completely removed, and the revised algorithm starts from a somewhat arbitrary initialization sampler coupled with a suitably adjusted importance weight. In short, the algorithm no longer relies on a prior distribution to initialize the process. This change takes advantage of cumulative knowledge on parameters' locations and dispersions to significantly speeds up optimization, but does not actually change the final optimization solution.

The revised algorithm can handle any prior distribution, but we chose not to impose any in order to avoid the flip side of relying on one's prior belief. A supposedly informative prior can be *subjective* to the point that it overpowers the information contained in the data, leading to a posterior outcome less reflective of the reality. This revision in the CRI implementation frees the estimation algorithm from needing an ad hoc prior belief to start the process. In effect, we set the prior to 1, i.e., a uniform or improper prior. 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.

II. Relax some NS parameters for the Chinese model estimation

We further revise the Chinese model by relaxing a couple of restrictions in the NS parameter estimation.[1] To see this simply, let $\beta(t,\tau;t_0)$ denote the coefficient of the two covariates — Intercept and DTD Level — bearing structural changes. Here, t is the time of performing prediction (e.g., now), τ is forecast horizon (measured in years) ranging from 0 to 59/12, and t_0 is the structural-break point, i.e., December 2004. The structural-break specification has the following form:

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

Further, each of $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$ has an NS function with four parameters, and for example,

$$\tilde{\gamma}(\tau;\rho_0,\rho_1,\rho_2,d) = \rho_0 + \rho_1\left[\frac{1-\exp(-\tau/d)}{\tau/d}\right] + \rho_2\left[\frac{1-\exp(-\tau/d)}{\tau/d} - \exp(-\tau/d)\right] \qquad (1)$$

Table 1 summarizes our new treatments on the four NS parameters $[\rho_0,\rho_1,\rho_2,d]$ concerning Intercept and DTD Level. The key points are as follow:

Table 1. Summary of Structural-Break Variables

The table presents the changes in the NS parameter estimation for Intercept and DTD level. The “1” in black denotes that the parameter was and is estimated. The “0” in black refers to setting the parameter value to 0 both previously and now. The “2” in red means the parameter was set to 0 previously but now it is estimated. Also, the range of d is relaxed to fall between 0 and infinity.

	ρ_0	ρ_1	ρ_2	d
Intercept $\tilde{\alpha}$	**2**	**1**	**1**	**1**
Intercept $\tilde{\beta}$	**1**	**2**	**1**	**1**
Intercept $\tilde{\gamma}$	**1**	**2**	**1**	**1**
DTD Level $\tilde{\alpha}$	**2**	**1**	**1**	**1**
DTD Level $\tilde{\beta}$	**0**	**1**	**1**	**1**
DTD Level $\tilde{\gamma}$	**0**	**1**	**1**	**1**

For the Intercept that is non-stochastic, we estimate the four parameters in all three NS functions. The reason is that as τ goes to infinity, the two terms in the brackets of the NS function (1) converge to zero, and ρ_0 should remain to capture the permanent effect on default and other exit.

[1] Interested readers are referred to Section 1.3.2 Structural break of the Technical Report or Version 2016 Update 1 Addendum 1*.

For the stochastic variables, except for $\tilde{\alpha}$ of the DTD level, we set ρ_0 to 0, because they should not be informative as the forecast horizon becomes infinitely large. Occasionally, one may experience an identification issue in distinguishing ρ_1 and ρ_2 when d is large. It is understandable because under a large d, the NS function is essentially a linear function of τ for up to our maximum prediction horizon of 5 years. Since only two parameters are needed to define a linear function, three parameters cannot be separately identified. When such an identification issue arises, we will set $\rho_2 = 0$. Note that the prior belief in essence provides the identification in the previous implementation.

The performance of the revised model with more flexibility and unnecessary restrictions removed is similar to the previous model. The accuracy ratio for 1-year, 2-year, and 5-year PDs are 67%, 65%, and 55%, receptively. Figure 1 plots the goodness of fit for the Chinese sample.

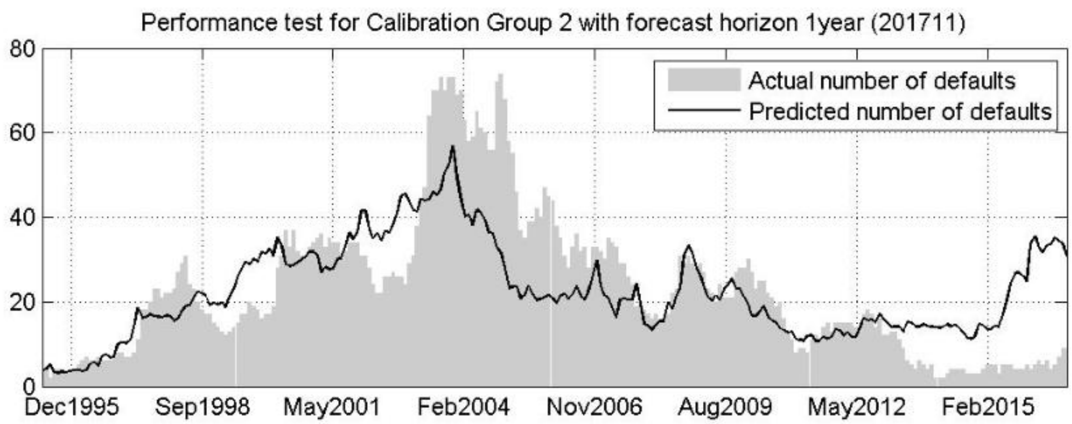

Figure 1. Goodness of Fit for China. The figure plots the performance test between the CRI 1-year PD estimated in December 2017 and the actual number of defaults.

References

Credit Research Initiative, 2016, "Version 2016 Update Addendum 1*: Revision to the parameter estimation on intercept and the DTD Level and their statistical inference for the Chinese sample," National University of Singapore

Credit Research Initiative, 2017, "NUS-RMI Credit Research Initiative Technical Report Version: 2017 update 1," National University of Singapore.

Credit Research Initiative, 2017, "Version 2017 Update Addendum 4: Changes in covariates in the CRI Probability of Default Model," National University of Singapore.

Duan, J.C., and A. Fulop, 2013, "Multiperiod Corporate Default Prediction with Partially-Conditioned Forward Intensity," National University of Singapore working paper.