Publication Date: 15 May 2013
Effective Date: 15 May 2013
Addendum 8 to the CRI Technical Report, (Version: 2012, Update 2)

Changes to Distance-to-Default (DTD) computation

This addendum updates the Technical Report (Version: 2012, Update 2) and reports the changes to the Distant-To-Default (DTD) computation. These changes have been implemented as of the DTD computation in May 2012 that was used for the probability of default (PD) released on 15 May 2013. The main aim of these changes is to greatly reduce the computational time of the DTD computation. Full detail on the DTD calculation is provided in this addendum for the sake of completeness, but not all details will be included in the next version of the technical report due to its length.

The DTD computation used in the CRI system is a two-stage procedure described in Section 3.2 of the Technical Report. A brief summary of the previous implementation follows:

In the first stage, the drift μ, the volatility σ and the fraction δ are computed by maximizing the log-likelihood function

$$\mathcal{L}(\mu,\sigma,\delta) = -\frac{n-1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\log(\sigma^2 h_t) - \sum_{t=2}^{n}\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\right) - \sum_{t=2}^{n}\log\left[N\left(\hat{d}_1\left(\hat{V}_t(\sigma,\delta),\sigma,\delta\right)\right)\right] - \frac{1}{2\sigma^2}\sum_{t=2}^{n}\frac{1}{h_t}\left[\log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\times\frac{A_{t-1}}{\hat{V}_{t-1}(\sigma,\delta)}\right) - \left(\mu - \frac{\sigma^2}{2}\right)h_t\right]^2 \quad (1)$$

over the region $\sigma \geq 0$ and $\delta_l \leq \delta \leq \delta_u$, where n is the number of days with observations of the equity value in the sample, h_t is the number of trading days as a fraction of the year between observations $t-1$ and t, $N(\cdot)$ is the standard normal cumulative distribution function, A_t is the book asset value, $\hat{V}_t$ is the implied asset value solved by using

$$E_t = V_t N(d_1) - e^{-r(T-t)} L N(d_2) \quad (2)$$

with

$$d_{1,2} = \frac{\log\left(\frac{V_t}{L_t}\right) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad (3)$$

and $\hat{d}_1$ is computed by using Equation (3) with $\hat{V}_t$. In Equation (2) and Equation (3), $T-t$ is the time to maturity set to be one year, r is the risk-free rate, E_t is the equity value, L_t is the debt level (also the default point) set to be sum of current liabilities, half of the long-term debt and the other liabilities multiplied by the fraction δ. This maximization for each firm is performed over all three variables, with lower and upper constraints $[\delta_l, \delta_u] = [0, 1]$ at month 1 and $[\delta_l, \delta_u] = [\max(0, \hat{\delta}_{n-1} -$

$0.05), \min(1, \hat{\delta}_{n-1} + 0.05)]$ at month n with $n \geq 2$, where $\hat{\delta}_{n-1}$ is the estimate of δ made at month $n-1$.

In the second stage, to impose great stability of the estimates of δ, all financial sector firms in the same economy are assumed to share the same estimate of δ, chosen to be the average of all its individual estimates, denoted by $\bar{\delta}$. The same is done for non-financial firms. Accordingly, with δ being fixed to be the sector average $\bar{\delta}$, the original maximization of $\mathcal{L}(\mu, \sigma, \delta)$ is reduced to a one-dimensional maximization in σ only, by using the fact that the estimates $\hat{\mu}$ and $\hat{\sigma}$ obtained from the maximization satisfy

$$\hat{\mu} = \frac{\hat{\sigma}^2}{2} + \frac{1}{\sum_{t=2}^{n} h_t} \log\left(\frac{\hat{V}_n(\hat{\sigma}, \hat{\delta})}{A_n} \times \frac{A_1}{\hat{V}_1(\hat{\sigma}, \hat{\delta})}\right). \tag{4}$$

Thus, this maximization is used to perform the estimates of σ for each firm. Finally, DTD is calculated by using the the estimates $\hat{\sigma}$ and $\hat{\delta}$ and the formula

$$\text{DTD}_t = \frac{1}{\sigma\sqrt{T-t}} \log\left(\frac{\hat{V}_t(\sigma, \delta)}{L_t}\right).$$

The maximization problems in the CRI system are solved by using the MATLAB Optimization Toolbox. The dimension reduction of the maximization described above in the second stage could help for both faster computation and better solution quality. However, as can be easily seen, the dimension reduction can also be applied to the maximization in the first stage. More precisely, by using the same relation (4), 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

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

The MATLAB function chosen for the maximization in the first stage is `fmincon`. Currently, the algorithm option of `fmincon` is changed from *interior-point* to *trust-region-reflective*, since the later one is more efficient for box constrained optimization problems in general. Meanwhile, the SQP algorithm is set to be the backup option of `fmincon` in case the output solutions of the trust-region-reflective algorithm do not achieve the required first order optimality. As the function $\tilde{\mathcal{L}}(\sigma, \delta)$ to be maximized is non-concave and `fmincon` only finds a local maximizer, the starting point passed to `fmincon` may affect the output solution. In view of this, for the first valid month of each company in which the maximization is performed, besides the old starting point, a new starting point, i.e., the overall average based on past experience, is also passed to `fmincon` for another trial, and the solution is chosen from the two resulting outputs by comparing their corresponding objective function values. In addition, for the other valid months of each firm, a warm-start strategy is applied to the maximization, i.e., using the solution of the previous valid month to be the starting point of the current month. This strategy not only accelerates the convergence of `fmincon` but also improves the stability of the estimates.

To use the trust-region-reflective algorithm, the gradient of the objective function should be supplied to `fmincon`. For notational simplicity, let

$$\theta_t(\sigma,\delta) := \log\left(\frac{\hat{V}_t(\sigma,\delta)}{A_t}\right).$$

By direct calculation, the gradient of $\tilde{\mathcal{L}}(\sigma,\delta)$ can be expressed as

$$\text{gradient} = \left(\frac{\partial\tilde{\mathcal{L}}}{\partial\sigma}, \frac{\partial\tilde{\mathcal{L}}}{\partial\delta}\right),$$

where

$$\begin{aligned}\frac{\partial\tilde{\mathcal{L}}}{\partial\sigma} &= \sum_{i=2}^{n}\frac{1}{\sigma} + \sum_{t=2}^{n}\frac{\partial\theta_t}{\partial\sigma} + \sum_{t=2}^{n}\frac{p(\hat{d}_1)}{N(\hat{d}_1)}\cdot\frac{\partial\hat{d}_1}{\partial\sigma} - \frac{1}{\sigma^3}\left[\sum_{t=2}^{n}\frac{1}{h_t}(\theta_t-\theta_{t-1})^2 - \frac{1}{\sum_{t=2}^{n}h_t}(\theta_n-\theta_1)^2\right] \\ &\quad + \frac{1}{\sigma^2}\left[\sum_{t=2}^{n}\frac{1}{h_t}(\theta_t-\theta_{t-1})\left(\frac{\partial\theta_t}{\partial\sigma}-\frac{\partial\theta_{t-1}}{\partial\sigma}\right) - \frac{1}{\sum_{t=2}^{n}h_t}(\theta_n-\theta_1)\left(\frac{\partial\theta_n}{\partial\sigma}-\frac{\partial\theta_1}{\partial\sigma}\right)\right],\end{aligned} \tag{5}$$

and

$$\begin{aligned}\frac{\partial\tilde{\mathcal{L}}}{\partial\delta} &= \sum_{t=2}^{n}\frac{\partial\theta_t}{\partial\delta} + \sum_{t=2}^{n}\frac{p(\hat{d}_1)}{N(\hat{d}_1)}\cdot\frac{\partial\hat{d}_1}{\partial\delta} \\ &\quad + \frac{1}{\sigma^2}\left[\sum_{t=2}^{n}\frac{1}{h_t}(\theta_t-\theta_{t-1})\left(\frac{\partial\theta_t}{\partial\delta}-\frac{\partial\theta_{t-1}}{\partial\delta}\right) - \frac{1}{\sum_{t=2}^{n}h_t}(\theta_n-\theta_1)\left(\frac{\partial\theta_n}{\partial\delta}-\frac{\partial\theta_1}{\partial\delta}\right)\right].\end{aligned}$$

Here, $p(\cdot)$ is the standard normal probability distribution function. The implicit differentiation of Equation (2) yields the explicit expressions of $\frac{\partial\theta_t}{\partial\sigma}$ and $\frac{\partial\theta_t}{\partial\delta}$ as

$$\frac{\partial\theta_t}{\partial\sigma} = \frac{1}{\hat{V}_t}\cdot\frac{\partial\hat{V}_t}{\partial\sigma} = \sqrt{T-t}\cdot\frac{\alpha_t}{\gamma_t} \qquad \text{and} \qquad \frac{\partial\theta_t}{\partial\delta} = \frac{1}{\hat{V}_t}\cdot\frac{\partial\hat{V}_t}{\partial\delta} = \frac{K_t}{L_t}\cdot\frac{\beta_t}{\gamma_t}, \tag{6}$$

where K_t is the firm's other liability, and

$$\begin{aligned}\alpha_t &:= p(\hat{d}_1)\hat{d}_2 - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}p(\hat{d}_2)\hat{d}_1, \\ \beta_t &:= p(\hat{d}_1) - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}\left[p(\hat{d}_2) - \sigma\sqrt{T-t}\,N(\hat{d}_2)\right], \\ \gamma_t &:= \sigma\sqrt{T-t}\,N(\hat{d}_1) + p(\hat{d}_1) - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}p(\hat{d}_2).\end{aligned} \tag{7}$$

Meanwhile, the differentiation of Equation (3) yields

$$\begin{aligned}\frac{\partial \hat{d}_1}{\partial \sigma} &= \frac{1}{\sigma}\left(\frac{1}{\sqrt{T-t}} \cdot \frac{\partial \theta_t}{\partial \sigma} - \hat{d}_2\right), \\ \frac{\partial \hat{d}_2}{\partial \sigma} &= \frac{1}{\sigma}\left(\frac{1}{\sqrt{T-t}} \cdot \frac{\partial \theta_t}{\partial \sigma} - \hat{d}_1\right), \\ \frac{\partial \hat{d}_1}{\partial \delta} = \frac{\partial \hat{d}_2}{\partial \delta} &= \frac{1}{\sigma\sqrt{T-t}}\left(\frac{\partial \theta_t}{\partial \delta} - \frac{K_t}{L_t}\right).\end{aligned} \tag{8}$$

To further improve the efficiency, the Hessian (matrix) of the objective function $\tilde{\mathcal{L}}(\sigma, \delta)$ is also supplied to `fmincon`, taking the form

$$\text{Hessian} = \begin{pmatrix} \frac{\partial^2 \tilde{\mathcal{L}}}{\partial \sigma^2} & \frac{\partial^2 \tilde{\mathcal{L}}}{\partial \sigma \partial \delta} \\ \frac{\partial^2 \tilde{\mathcal{L}}}{\partial \delta \partial \sigma} & \frac{\partial^2 \tilde{\mathcal{L}}}{\partial \delta^2} \end{pmatrix},$$

where

$$\begin{aligned}\frac{\partial^2 \tilde{\mathcal{L}}}{\partial \sigma^2} = &-\sum_{t=2}^{n} \frac{1}{\sigma^2} + \sum_{t=2}^{n} \frac{\partial^2 \theta_t}{\partial \sigma^2} + \sum_{t=2}^{n} \frac{p(\hat{d}_1)}{N(\hat{d}_1)}\left[\frac{\partial^2 \hat{d}_1}{\partial \sigma^2} - \left(\hat{d}_1 + \frac{p(\hat{d}_1)}{N(\hat{d}_1)}\right)\left(\frac{\partial \hat{d}_1}{\partial \sigma}\right)^2\right] \\ &+ \frac{3}{\sigma^4}\left[\sum_{t=2}^{n} \frac{1}{h_t}(\theta_t - \theta_{t-1})^2 - \frac{1}{\sum_{t=2}^{n} h_t}(\theta_n - \theta_1)^2\right] \\ &- \frac{4}{\sigma^3}\left[\sum_{t=2}^{n} \frac{1}{h_t}(\theta_t - \theta_{t-1})\left(\frac{\partial \theta_t}{\partial \sigma} - \frac{\partial \theta_{t-1}}{\partial \sigma}\right) - \frac{1}{\sum_{t=2}^{n} h_t}(\theta_n - \theta_1)\left(\frac{\partial \theta_n}{\partial \sigma} - \frac{\partial \theta_1}{\partial \sigma}\right)\right] \\ &+ \frac{1}{\sigma^2}\left\{\sum_{t=2}^{n} \frac{1}{h_t}\left[\left(\frac{\partial \theta_t}{\partial \sigma} - \frac{\partial \theta_{t-1}}{\partial \sigma}\right)^2 + (\theta_t - \theta_{t-1})\left(\frac{\partial^2 \theta_t}{\partial \sigma^2} - \frac{\partial^2 \theta_{t-1}}{\partial \sigma^2}\right)\right]\right. \\ &\left.\quad - \frac{1}{\sum_{t=2}^{n} h_t}\left[\left(\frac{\partial \theta_n}{\partial \sigma} - \frac{\partial \theta_1}{\partial \sigma}\right)^2 + (\theta_n - \theta_1)\left(\frac{\partial^2 \theta_n}{\partial \sigma^2} - \frac{\partial^2 \theta_1}{\partial \sigma^2}\right)\right]\right\},\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial^2 \tilde{\mathcal{L}}}{\partial \delta^2} = &\sum_{t=2}^{n} \frac{\partial^2 \theta_t}{\partial \delta^2} + \sum_{t=2}^{n} \frac{p(\hat{d}_1)}{N(\hat{d}_1)}\left[\frac{\partial^2 \hat{d}_1}{\partial \delta^2} - \left(\hat{d}_1 + \frac{p(\hat{d}_1)}{N(\hat{d}_1)}\right)\left(\frac{\partial \hat{d}_1}{\partial \delta}\right)^2\right] \\ &+ \frac{1}{\sigma^2}\left\{\sum_{t=2}^{n} \frac{1}{h_t}\left[\left(\frac{\partial \theta_t}{\partial \delta} - \frac{\partial \theta_{t-1}}{\partial \delta}\right)^2 + (\theta_t - \theta_{t-1})\left(\frac{\partial^2 \theta_t}{\partial \delta^2} - \frac{\partial^2 \theta_{t-1}}{\partial \delta^2}\right)\right]\right. \\ &\left.\quad - \frac{1}{\sum_{t=2}^{n} h_t}\left[\left(\frac{\partial \theta_n}{\partial \delta} - \frac{\partial \theta_1}{\partial \delta}\right)^2 + (\theta_n - \theta_1)\left(\frac{\partial^2 \theta_n}{\partial \delta^2} - \frac{\partial^2 \theta_1}{\partial \delta^2}\right)\right]\right\},\end{aligned}$$

and

$$
\begin{aligned}
\frac{\partial^2\tilde{\mathcal{L}}}{\partial\sigma\partial\delta} = \frac{\partial^2\tilde{\mathcal{L}}}{\partial\delta\partial\sigma} &= \sum_{t=2}^{n}\frac{\partial^2\theta_t}{\partial\sigma\partial\delta} + \sum_{t=2}^{n}\frac{p(\hat{d}_1)}{N(\hat{d}_1)}\left[\frac{\partial^2\hat{d}_1}{\partial\sigma\partial\delta} - \left(\hat{d}_1 + \frac{p(\hat{d}_1)}{N(\hat{d}_1)}\right)\cdot\frac{\partial\hat{d}_1}{\partial\sigma}\cdot\frac{\partial\hat{d}_1}{\partial\delta}\right] \\
&\quad - \frac{2}{\sigma^3}\left[\sum_{t=2}^{n}\frac{1}{h_t}(\theta_t - \theta_{t-1})\left(\frac{\partial\theta_t}{\partial\delta} - \frac{\partial\theta_{t-1}}{\partial\delta}\right) - \frac{1}{\sum_{t=2}^{n}h_t}(\theta_n - \theta_1)\left(\frac{\partial\theta_n}{\partial\delta} - \frac{\partial\theta_1}{\partial\delta}\right)\right] \\
&\quad + \frac{1}{\sigma^2}\left\{\sum_{t=2}^{n}\frac{1}{h_t}\left[\left(\frac{\partial\theta_t}{\partial\sigma} - \frac{\partial\theta_{t-1}}{\partial\sigma}\right)\left(\frac{\partial\theta_t}{\partial\delta} - \frac{\partial\theta_{t-1}}{\partial\delta}\right) + (\theta_t - \theta_{t-1})\left(\frac{\partial^2\theta_t}{\partial\sigma\partial\delta} - \frac{\partial^2\theta_{t-1}}{\partial\sigma\partial\delta}\right)\right]\right. \\
&\quad\quad \left. - \frac{1}{\sum_{t=2}^{n}h_t}\left[\left(\frac{\partial\theta_n}{\partial\sigma} - \frac{\partial\theta_1}{\partial\sigma}\right)\left(\frac{\partial\theta_n}{\partial\delta} - \frac{\partial\theta_1}{\partial\delta}\right) + (\theta_n - \theta_1)\left(\frac{\partial^2\theta_n}{\partial\sigma\partial\delta} - \frac{\partial^2\theta_1}{\partial\sigma\partial\delta}\right)\right]\right\}.
\end{aligned}
$$

The differentiation of Equation (6) yields the explicit expressions

$$
\begin{aligned}
\frac{\partial^2\theta_t}{\partial\sigma^2} &= -\frac{1}{\hat{V}_t^2}\left(\frac{\partial\hat{V}_t}{\partial\sigma}\right)^2 + \frac{1}{\hat{V}_t}\cdot\frac{\partial^2\hat{V}_t}{\partial\sigma^2} = \frac{\sqrt{T-t}}{\gamma_t^2}\left(\frac{\partial\alpha_t}{\partial\sigma}\gamma_t - \alpha_t\frac{\partial\gamma_t}{\partial\sigma}\right), \\
\frac{\partial^2\theta_t}{\partial\delta^2} &= -\frac{1}{\hat{V}_t^2}\left(\frac{\partial\hat{V}_t}{\partial\delta}\right)^2 + \frac{1}{\hat{V}_t}\cdot\frac{\partial^2\hat{V}_t}{\partial\delta^2} = -\frac{K_t^2}{L_t^2}\cdot\frac{\beta_t}{\gamma_t} + \frac{K_t}{L_t}\cdot\frac{1}{\gamma_t^2}\left(\frac{\partial\beta_t}{\partial\delta}\gamma_t - \beta_t\frac{\partial\gamma_t}{\partial\delta}\right), \\
\frac{\partial^2\theta_t}{\partial\sigma\partial\delta} = \frac{\partial^2\theta_t}{\partial\delta\partial\sigma} &= -\frac{1}{\hat{V}_t^2}\cdot\frac{\partial\hat{V}_t}{\partial\sigma}\cdot\frac{\partial\hat{V}_t}{\partial\delta} + \frac{1}{\hat{V}_t}\cdot\frac{\partial^2\hat{V}_t}{\partial\sigma\partial\delta} = \frac{K_t}{L_t}\cdot\frac{1}{\gamma_t^2}\left(\frac{\partial\beta_t}{\partial\sigma}\gamma_t - \beta_t\frac{\partial\gamma_t}{\partial\sigma}\right),
\end{aligned}
$$

where α_t, β_t and γ_t are given by Equation (7), and

$$
\begin{aligned}
\frac{\partial\alpha_t}{\partial\sigma} &= p(\hat{d}_1)\left(\frac{\partial\hat{d}_2}{\partial\sigma} - \hat{d}_1\hat{d}_2\frac{\partial\hat{d}_1}{\partial\sigma}\right) - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}p(\hat{d}_2)\left(\frac{\partial\hat{d}_1}{\partial\sigma} - \hat{d}_1\frac{\partial\theta_t}{\partial\sigma} - \hat{d}_1\hat{d}_2\frac{\partial\hat{d}_2}{\partial\sigma}\right), \\
\frac{\partial\beta_t}{\partial\sigma} &= -p(\hat{d}_1)\hat{d}_1\frac{\partial\hat{d}_1}{\partial\sigma} - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}\left[\left(p(\hat{d}_2) - \sigma\sqrt{T-t}\,N(\hat{d}_2)\right)\left(-\frac{\partial\theta_t}{\partial\sigma}\right)\right. \\
&\quad\quad \left. - p(\hat{d}_2)\left(\hat{d}_2 + \sigma\sqrt{T-t}\right)\frac{\partial\hat{d}_2}{\partial\sigma} - \sqrt{T-t}\,N(\hat{d}_2)\right], \\
\frac{\partial\beta_t}{\partial\delta} &= -p(\hat{d}_1)\hat{d}_1\frac{\partial\hat{d}_1}{\partial\delta} - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}\left[\left(p(\hat{d}_2) - \sigma\sqrt{T-t}\,N(\hat{d}_2)\right)\left(\frac{K_t}{L_t} - \frac{\partial\theta_t}{\partial\delta}\right)\right. \\
&\quad\quad \left. - p(\hat{d}_2)\left(\hat{d}_2 + \sigma\sqrt{T-t}\right)\frac{\partial\hat{d}_2}{\partial\delta}\right], \\
\frac{\partial\gamma_t}{\partial\sigma} &= \sqrt{T-t}\,N(\hat{d}_1) + p(\hat{d}_1)\left(\sigma\sqrt{T-t} - \hat{d}_1\right)\frac{\partial\hat{d}_1}{\partial\sigma} - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}p(\hat{d}_2)\left(-\frac{\partial\theta_t}{\partial\sigma} - \hat{d}_2\frac{\partial\hat{d}_2}{\partial\sigma}\right), \\
\frac{\partial\gamma_t}{\partial\delta} &= p(\hat{d}_1)\left(\sigma\sqrt{T-t} - \hat{d}_1\right)\frac{\partial\hat{d}_1}{\partial\delta} - e^{-r(T-t)}\frac{L_t}{\hat{V}_t}p(\hat{d}_2)\left(\frac{K_t}{L_t} - \frac{\partial\theta_t}{\partial\delta} - \hat{d}_2\frac{\partial\hat{d}_2}{\partial\delta}\right).
\end{aligned}
$$

Meanwhile, the differentiation of Equation (8) further yields

$$\frac{\partial^2 \hat{d}_1}{\partial \sigma^2} = \frac{1}{\sigma}\left(\frac{1}{\sqrt{T-t}}\frac{\partial^2 \theta_t}{\partial \sigma^2} - \frac{\partial \hat{d}_1}{\partial \sigma} - \frac{\partial \hat{d}_2}{\partial \sigma}\right),$$

$$\frac{\partial^2 \hat{d}_1}{\partial \delta^2} = \frac{1}{\sigma\sqrt{T-t}}\left(\frac{\partial^2 \theta_t}{\partial \delta^2} + \frac{K_t^2}{L_t^2}\right),$$

$$\frac{\partial^2 \hat{d}_1}{\partial \sigma \partial \delta} = \frac{\partial^2 \hat{d}_1}{\partial \delta \partial \sigma} = \frac{1}{\sigma}\left(\frac{1}{\sqrt{T-t}}\frac{\partial^2 \theta_t}{\partial \sigma \partial \delta} - \frac{\partial \hat{d}_2}{\partial \delta}\right).$$

The main computational cost in each iteration of the maximization lies in the evaluation of the implied asset value $\hat{V}_t$ by solving the nonlinear system (2). As can be seen from above, in each iteration, after the evaluation of the objective function $\tilde{\mathcal{L}}(\sigma, \delta)$, the additional evaluations of the gradient and the Hessian use the same $\hat{V}_t$ and thus need only relatively little time. However, if the gradient and the Hessian were not supplied to `fmincon`, as before in the previous use of `fmincon`, finite-difference approximations would be used instead so that more nonlinear systems would need to be solved. Therefore, passing the gradient and the Hessian to `fmincon` can significantly reduce the computational time of the maximization.

To evaluate $\hat{V}_t$, Newton's method is applied to solve the nonlinear system (2), while bisection search is set to be the backup when the Newton's method fails to converge. It is well-known that the efficiency of Newton's method relies on a well-chosen starting point. Due to this, the previous treatment is first to find an (multi-dimensional) interval that contains the solution and then to apply Newton's method with the starting point set to be the center of the interval. However, using this treatment, apart from the computational cost of finding a suitable interval, Newton's method can fail to converge from time to time so that bisection search has to be further applied, resulting in slow convergence. Currently, the previous treatment is changed to be a direct implementation of Newton's method with a warm-start strategy, i.e., using the implied asset value estimated in the previous month to be the starting point in the current month. This warm-start strategy makes Newton's method seldom fail and also accelerates the convergence of Newton's method. Therefore, the computational time of the maximization can be further greatly reduced.

The MATLAB function chosen for the maximization in the second stage is also changed, from `fminbnd` to `fmincon` with the algorithm option *trust-region-reflective*. Meanwhile, `fminbnd` is set to be the backup in case the output solutions of `fmincon` with the trust-region-reflective algorithm do not achieve the required first order optimality. All the ways described above to reduce the computational time in the first stage are also done in the second stage. In particular, the gradient and the Hessian of the function to be maximized, i.e., $\bar{\mathcal{L}}(\sigma) := \tilde{\mathcal{L}}(\sigma, \hat{\delta})$, takes the form

$$\text{gradient} = \frac{\partial \bar{\mathcal{L}}}{\partial \sigma} = \left.\frac{\partial \tilde{\mathcal{L}}}{\partial \sigma}\right|_{\delta=\bar{\delta}},$$

and

$$\text{Hessian} = \frac{\partial^2 \bar{\mathcal{L}}}{\partial \sigma^2} = \left.\frac{\partial \tilde{\mathcal{L}}}{\partial \sigma^2}\right|_{\delta=\bar{\delta}},$$

where $\frac{\partial \tilde{\mathcal{L}}}{\partial \sigma}$ and $\frac{\partial \tilde{\mathcal{L}}}{\partial \sigma^2}$ are given by Equation (5) and Equation (9) respectively.

To summarize, owing to the changes, the total computational time of the DTD calculation can be reduced from several months to about 70 hours on a single PC for all firms over the full history. In the current CRI system, by using a computational grid administered by the NUS Computer Center, the total computational time would take only about 3.5 hours.[1]

[1]The previous DTD computation used a different grid which was about 2.5 times faster than the current one, and the computational time was less than one day, as reported in Technical Report (Version: 2012, Update 2).