Data files:

1. **Liborswapdata.csv**

   Format: date(excelserial number), (1,2,3,6,9,12)-month Libor, (2,3,4,5,7,10,15,20,30)-year swap rates

   Table 1 summary statistics can be directly computed from these rates. The data are sampled weekly.

2. **Model3outputs.csv**

   Format: date, (x1,x2,x3)(3 factors), fitted values on the 15-libor and swap rates in the same order as the previous data

3. **Model15output.csv**

   Format: date, (x1,x2,x3, $\cdots$,x15)(15 factors), fitted values on the 15-libor and swap rates in the same order as the previous data

4. **osXhat15.csv**

   Format: date, (x1,x2,x3,$\cdots$,x15)

   Fitted factor values $\widehat{X}_t$ for out of sample testing periods.

   The values $\widehat{X}_t$ at each date t are generated from a different set of parameter estimates. The parameters are estimated using data up to that date t.

5. **osXPred15h1.csv**

   Format: date, (x1,x2,x3,$\cdots$,x15)

   Forecasted factor values $\overline{X}_{t+h}$ for out of sample testing periods from week 158 (1998/1/7) forward, with $h = 1$ week.

   For ease of comparison, the file length is kept the same as the other files, but filled with NaNs for the first 157 weeks.

6. **osXPred15h2.csv**

   Format: date, (x1,x2,x3,$\cdots$,x15)

   Forecasted factor values $\overline{X}_{t+h}$ for out of sample testing periods, with $h = 2$ weeks.

7. **osXPred15h3.csv**

   Format: date, (x1,x2,x3,$\cdots$,x15)

   Forecasted factor values $\overline{X}_{t+h}$ for out of sample testing periods, with $h = 3$ weeks.

8. **osXPred15h4.csv**

   Format: date, (x1,x2,x3,$\cdots$,x15)

   Forecasted factor values $\overline{X}_{t+h}$ for out of sample testing periods, with $h = 4$ week.

9. **osparameters15.csv**

   Format: date, parameters

   The estimated parameter values at each date t using data up to that date for the 15-factor model.

   Let p denote the list of parameters. They correspond to the model parameters in the paper as follows:

   $$\kappa_1 = e^{p_1}, \quad \sigma_r = e^{p_2}, \quad \theta_r^{\mathbb{Q}} = e^{p_3}, \quad b = e^{e^{p_4}}, \quad \theta_r = e^{p_5}, \quad \sigma_e^2 = e^{p_6}$$

Output Statistics:

1. **Table 1**

   Summary statistics can be directly computed from the libor and swap rates from **Liborswapdata.csv**.

2. **Table 3**

   The pricing error statistics in Table 3 can be replicated from data in three files. The pricing errors are defined as Error= (rates - fitted values)*100 (in basis points)

   The rates are from **Liborswapdata.csv**. The fitted values are from **Model3outputs.csv** for the three-factor model and **Model15outputs.csv** for the 15-factor model.

   The reported statistics in the paper are computed using the pricing errors from the 4th week forward. The first three observations are discarded from the statistics as they can be heavily affected by the initial state assumptions.

3. **Table 4B and 4C**

   The in-sample predictive variation in Table 4B and 4C, can be replicated from the extracted factors and the parameter estimates.

   At each date t, the factors in the data files can be regarded as $\widehat{X}_t$. Its predicted values h-years ahead are computed as

   $$\overline{X}_{t+h} = A + \Phi\widehat{X}_t,$$

   where A and Φ are functions of the estimated model parameters and forecasting horizons:

   $$\Phi = \exp(-\kappa h), \quad A = (I - \Phi)\theta,$$

   where κ is the mean-reversion matrix with the cascade structure, θ is a vector of the long-run mean θ_r.

   The forecasted values of the interest rates are directly computed based on the forecasted values of the factors $\overline{X}_{t+h}$.

4. **Table 5B**

   The out-of-sample predictive variation is computed in a more complex way:

   (a) At each date t, we use the interest rates data up to time t to re-estimate the model parameters. With the re-estimated model parameters, we run through the sample to obtain the fitted factor value at date t, $\widehat{X}_t$. — The difference between $\widehat{X}_t$ here for the out-of-sample statistics and $\widehat{X}_t$ for Table 4 in-sample statistics is that the extracted factors in the in-sample statistics are based on one-set of parameter estimates whereas for the out-of-sample statistics, the factors at each date during the out-of-sample period is based on a different set of parameter estimates.
   The fitted factors $\widehat{X}_t$ for the out-of-sample statistics for the 15-factor model are in **osXhat15.csv**.

   (b) The forecasting steps are the same as for the in-sample analysis.

   For ease of replication, we have also saved the predicted factor values at the reported four horizons, $\overline{X}_{t+h}$ with $h =$1,2,3,4 weeks. They are in files **osXPred15h1.csv**, **osXPred15h2.csv**, **osXPred15h3.csv**, **osXPred15h4.csv**, respectively.

   Furthermore, we have also provided the parameter estimates obtained at different dates in file **osparameters15.csv**. The estimates in the last date represent the full-sample estimates.