A BAYESIAN MULTIVARIATE FUNCTIONAL DYNAMIC LINEAR MODEL

Daniel R. Kowal

May 20, 2017

Cornell University and Rice University

Joint work with David S. Matteson and David Ruppert

INSTALL THE PACKAGE FDLM

Install and load the package `FDLM` from GitHub:

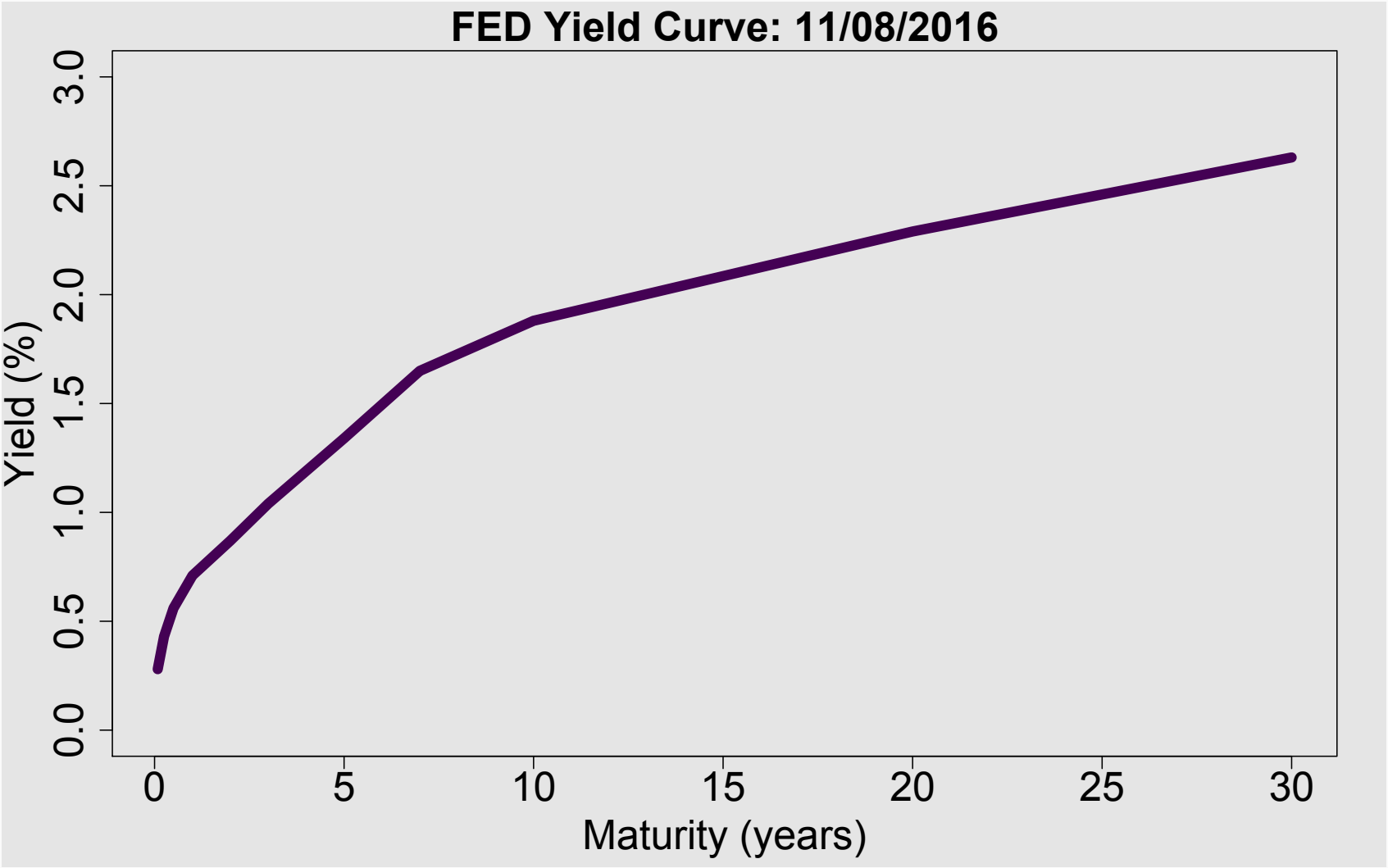
```
devtools::install_github("drkowal/FDLM")
```

```
library(FDLM)
```

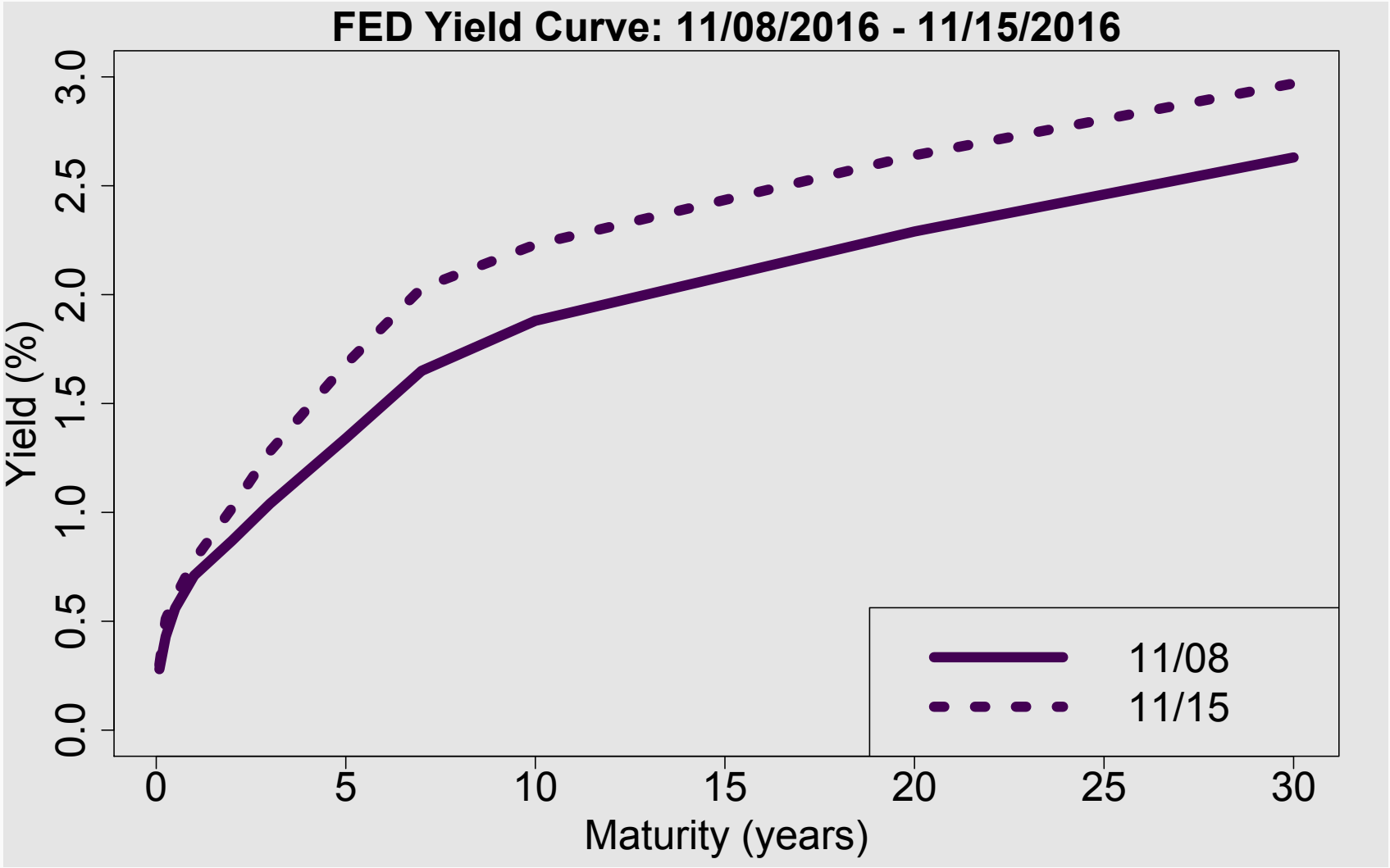
MULTIECONOMY YIELD CURVES:

A MULTIVARIATE TIME SERIES OF FUNCTIONAL DATA

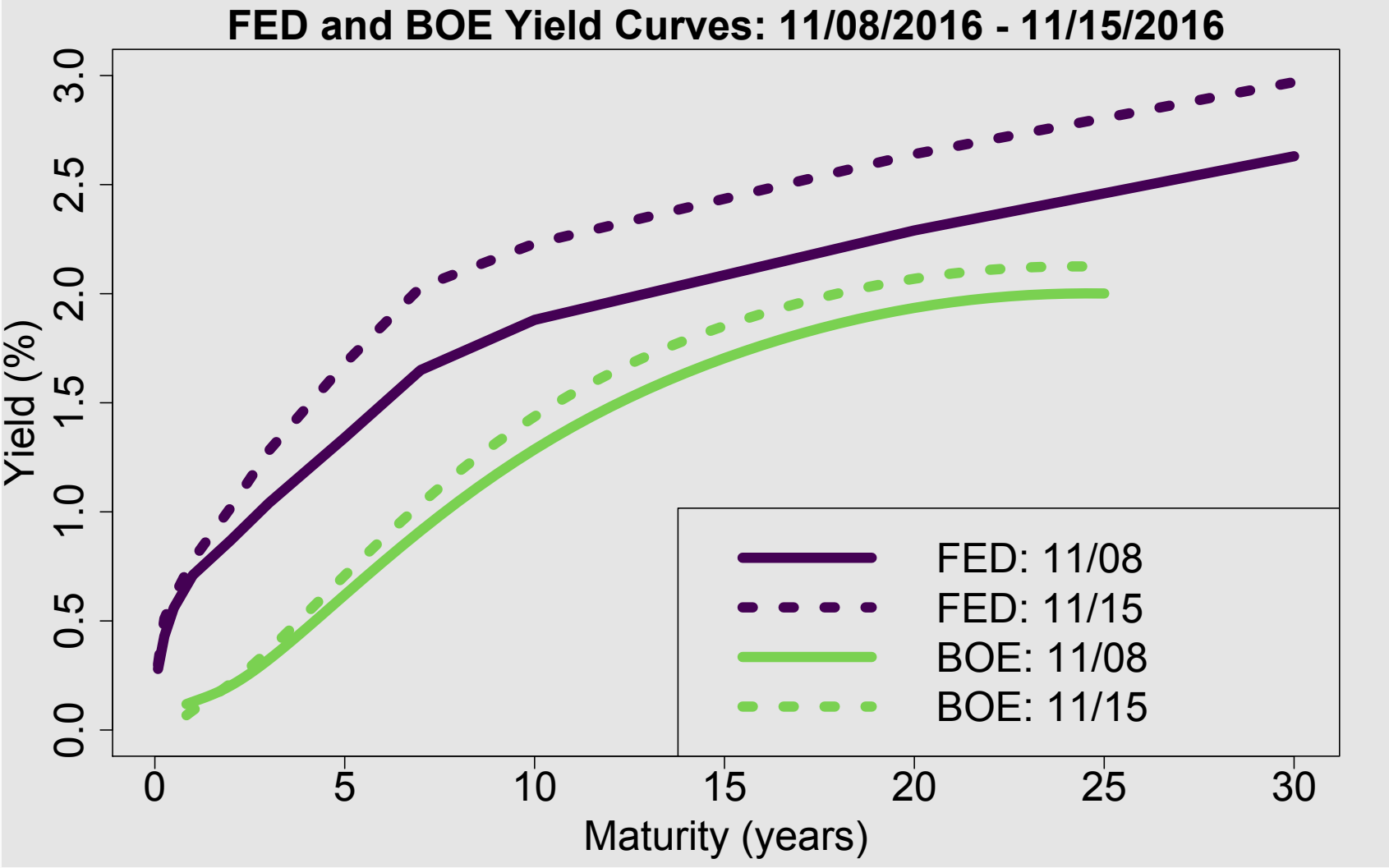
FUNCTIONAL: $y(\tau)$



FUNCTIONAL (τ) + TIME SERIES (t): $y_t(\tau)$



FUNCTIONAL (τ) + TIME SERIES (t) + MULTIVARIATE (c): $y_t^{(c)}(\tau)$



GENERAL SETTING: MULTIVARIATE TIME SERIES OF FUNCTIONAL DATA

Dominant structural features in the data:

1. Functional
2. Time-dependence: time-ordered observations
3. Contemporaneous dependence: multivariate observations

A MULTIVARIATE FUNCTIONAL DYNAMIC LINEAR MODEL

A FUNCTIONAL DYNAMIC LINEAR MODEL (FDLM)

Functional Dynamic Linear Model (FDLM) [Kowal et al., 2016]:

$$y_t(\tau) = \sum_{k=1}^{K} f_k(\tau)\beta_{k,t} + \epsilon_t(\tau), \quad \tau \in \mathcal{T}, \quad t = 1, \ldots, T$$

Decompose a functional time series y_t into

- *functional component* $\{f_k(\tau)\}_{k=1}^{K}$
- *time series component* $\left\{\beta_{k,t}\right\}_{k=1}^{K}$

We model $\boldsymbol{\beta}_t = (\beta_{1,t}, \ldots, \beta_{K,t})'$ as the state vector in a DLM

FDLM AS A FACTOR ANALYSIS

Functional Dynamic Linear Model (FDLM) [Kowal et al., 2016]:

$$y_t(\tau) = \sum_{k=1}^{K} f_k(\tau)\beta_{k,t} + \epsilon_t(\tau), \quad \tau \in \mathcal{T}, \quad t = 1, \ldots, T$$

Dynamic functional factor model:

- $\{f_k\}$ are *factor loading curves*
- $\{\beta_{k,t}\}$ are dynamic *factors*
- $K =$ number of factors

FDLM AS A DYNAMIC LINEAR MODEL

Suppose we observe $\boldsymbol{y}_t = (y_t(\tau_1), \ldots, y_t(\tau_M))'$

Functional Dynamic Linear Model (FDLM):

$$\begin{cases} \underbrace{\boldsymbol{y}_t}_{M\times 1} = \underbrace{\boldsymbol{F}}_{M\times K} \underbrace{\beta_t}_{K\times 1} + \underbrace{\epsilon_t}_{M\times 1}, & \epsilon_t \overset{indep}{\sim} N(\mathbf{o}, \sigma^2_{\epsilon_t} \boldsymbol{I}_M) \\ \\ \beta_t = \boldsymbol{G}\beta_{t-1} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t \overset{indep}{\sim} N(\mathbf{o}, \boldsymbol{W}_t) \end{cases}$$

where $\boldsymbol{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_K)$, and $\boldsymbol{f}_k = (f_k(\tau_1), \ldots, f_k(\tau_M))'$

Extensions for covariates, stochastic volatility, change points

Functional Dynamic Linear Model (FDLM):

$$\begin{cases} \underbrace{\boldsymbol{y}_t}_{M\times 1} = \underbrace{\boldsymbol{F}}_{M\times K}\underbrace{\boldsymbol{\beta}_t}_{K\times 1} + \underbrace{\boldsymbol{\epsilon}_t}_{M\times 1}, & \boldsymbol{\epsilon}_t \overset{indep}{\sim} N(\mathbf{o}, \sigma^2_{\epsilon_t}\boldsymbol{I}_M) \\ \\ \boldsymbol{\beta}_t = \boldsymbol{G}\boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t \overset{indep}{\sim} N(\mathbf{o}, \boldsymbol{W}_t) \end{cases}$$

For `R` implementation, use a state space package, such as `KFAS` [Helske, 2017]:

```
SSModel(yt ~ -1 + SSMcustom(Z = F,
                            T = G,
                            Q = Wt), H = sigma_et^2)
```

ESTIMATING THE FACTOR LOADING CURVES f_k

Nonparametric regression model for each f_k:

$$f_k(\tau) = \phi'(\tau)\boldsymbol{d}_k$$

- $\phi(\cdot)$ *known* vector of basis functions (splines)
- $\boldsymbol{d}_k$ *unknown* vector of basis coefficients
- Include roughness penalty via the prior

$$\boldsymbol{d}_k \sim N\left(\mathbf{o}, \lambda_k^{-1}\boldsymbol{\Omega}^{-1}\right)$$

Flexible, computationally efficient, and smooth

MODEL IMPLEMENTATION: MCMC SAMPLER

FDLM SAMPLING ALGORITHM

1. Sample the factors, $\{\beta_{k,t}\}$, using code from `KFAS`:

```
Beta = fdlm_factor(...)
```

2. Sample the observation and evolution error variances:

```
sigma_e = 1/sqrt(rgamma(...)
Wt = sample_Wt(...)
```

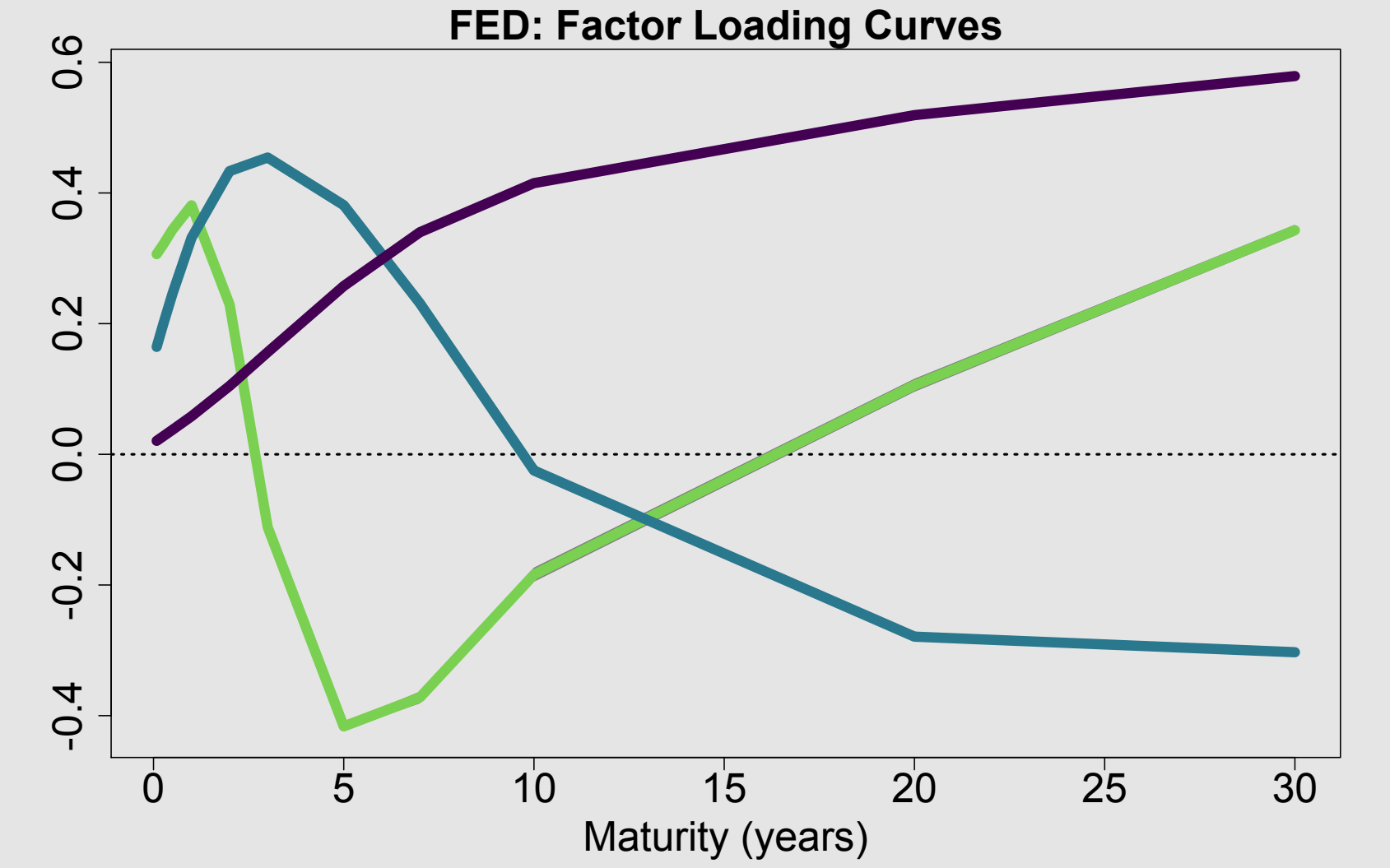
3. Sample the factor loading curves, $\{f_k\}$

```
D = sampleFLC(...)        # basis coefficients
```

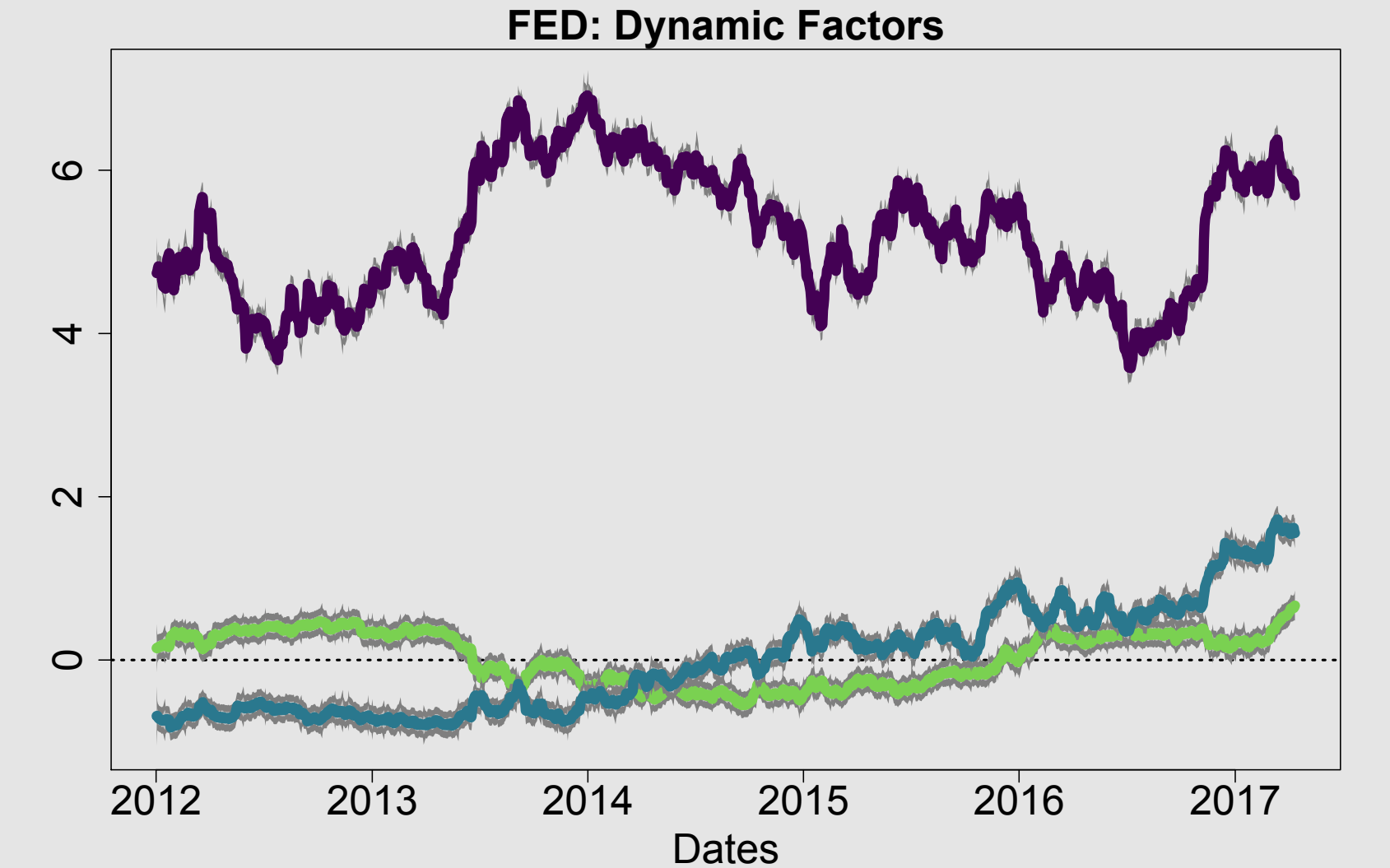
FDLM CODE EXAMPLE: fdlm()

```
# Read in the FED yield curve data (Y, tau, dates):
data("US_Yields")
#data("UK_Yields")

# Restrict to dates since 2012:
Y = Y[which(dates > as.Date("2012-01-01")),];

# Run the MCMC:
mcmc_output = fdlm(Y, tau,
                   K = 3,
                   nsims = 2500, burnin = 2500,
                   mcmc_params = list("beta", "fk",
                                      "Yhat", "sigma_e"))
```

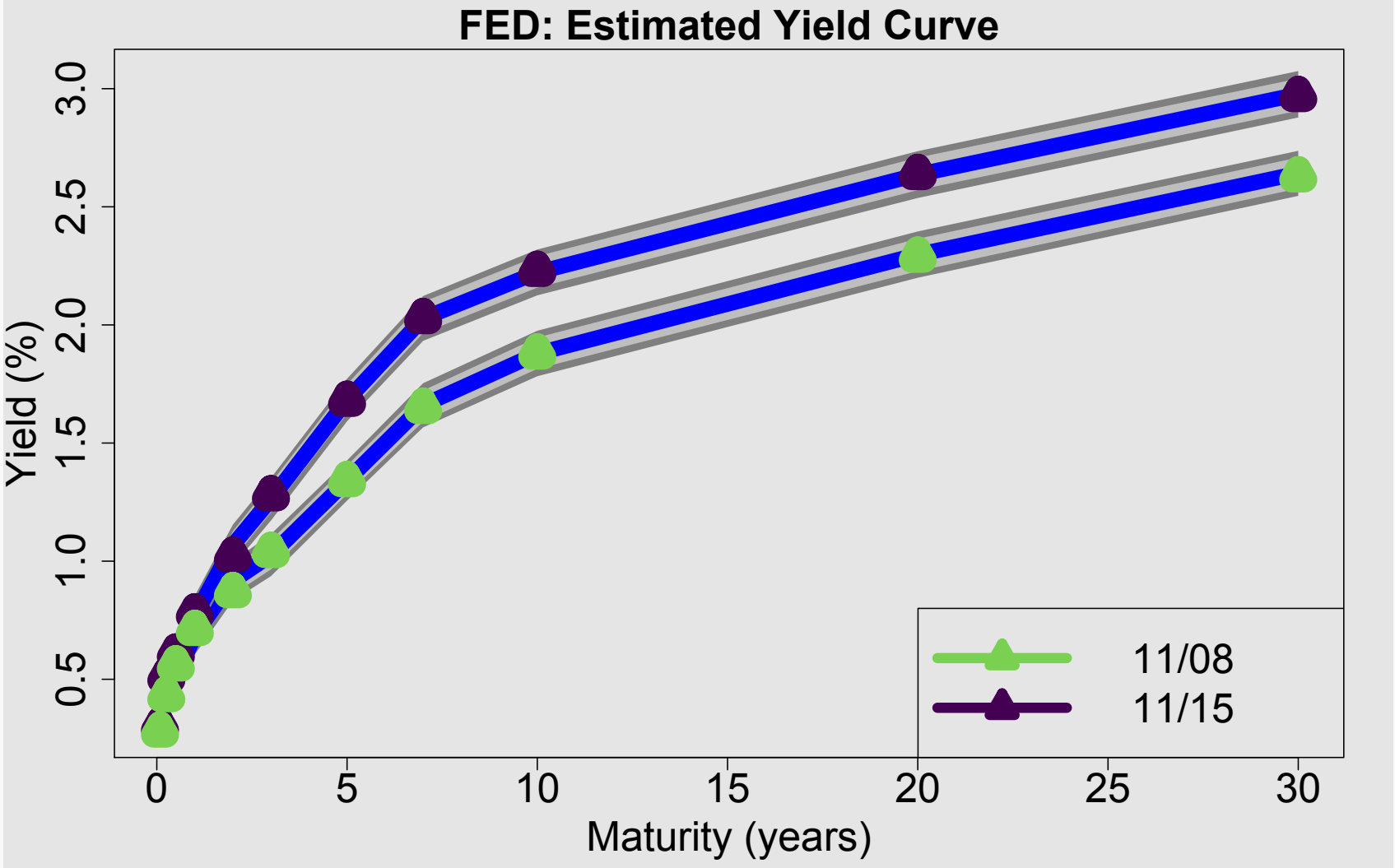
FACTOR LOADING CURVES: FED



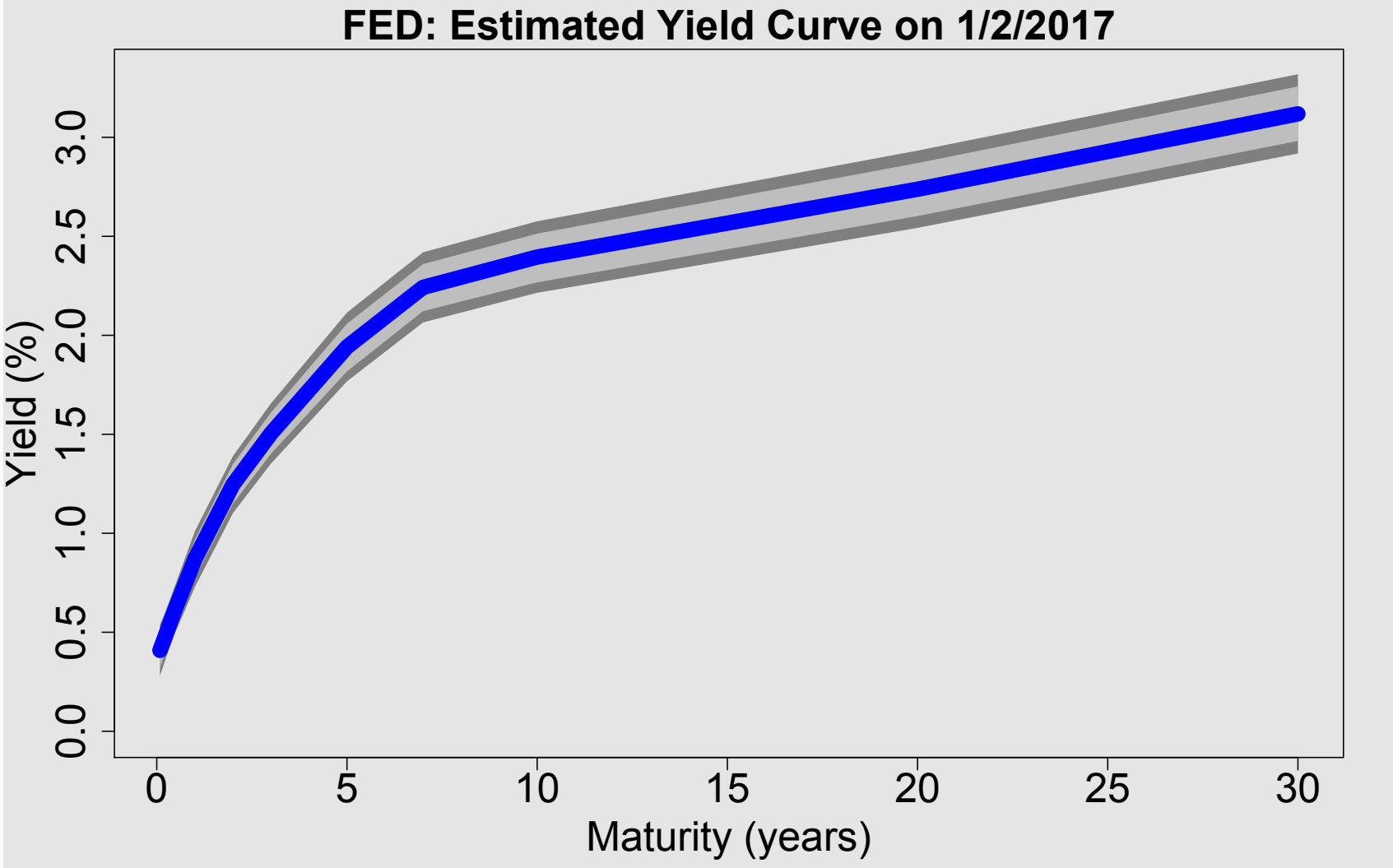
DYNAMIC FACTORS: FED



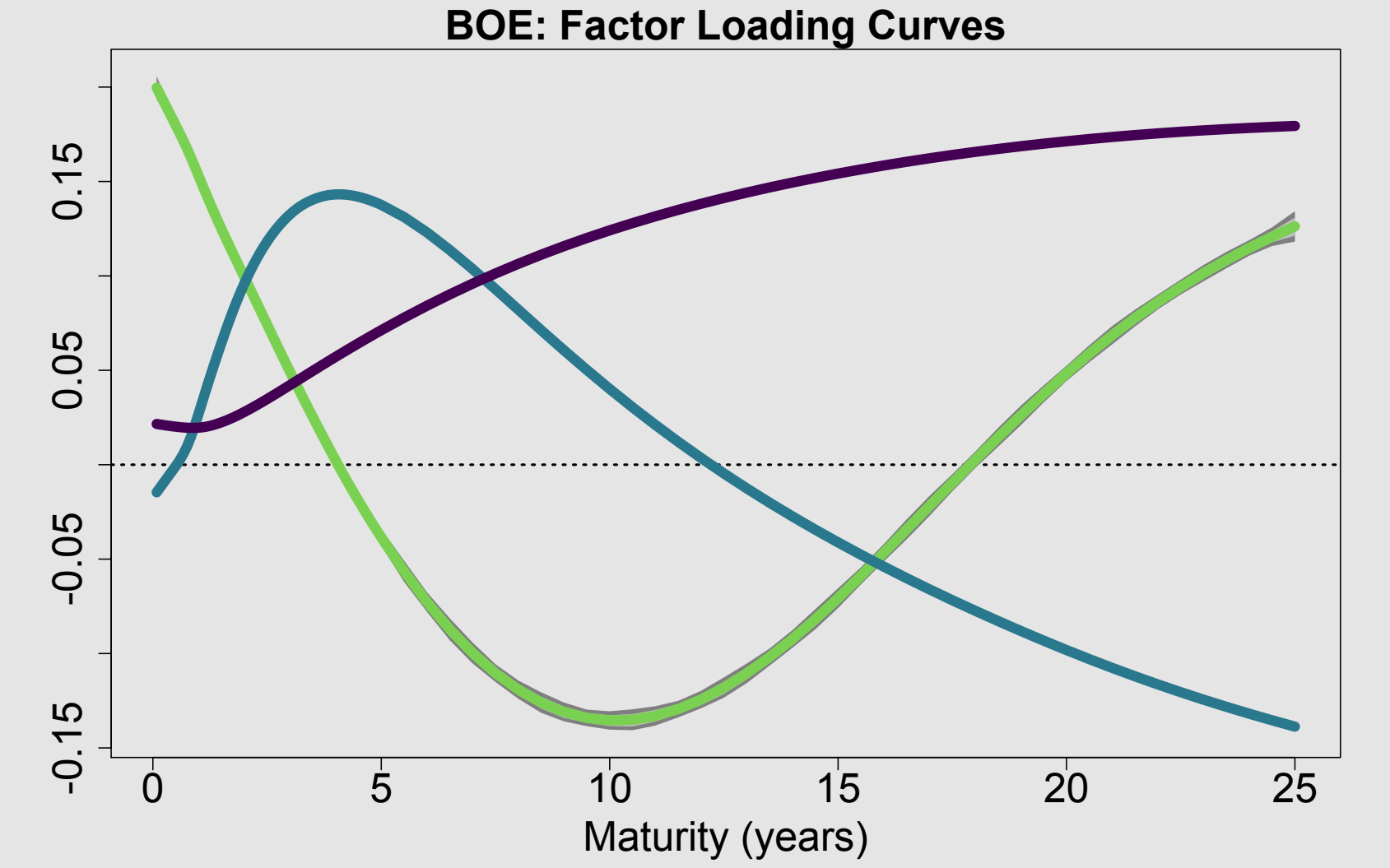
FITTED YIELD CURVE: FED



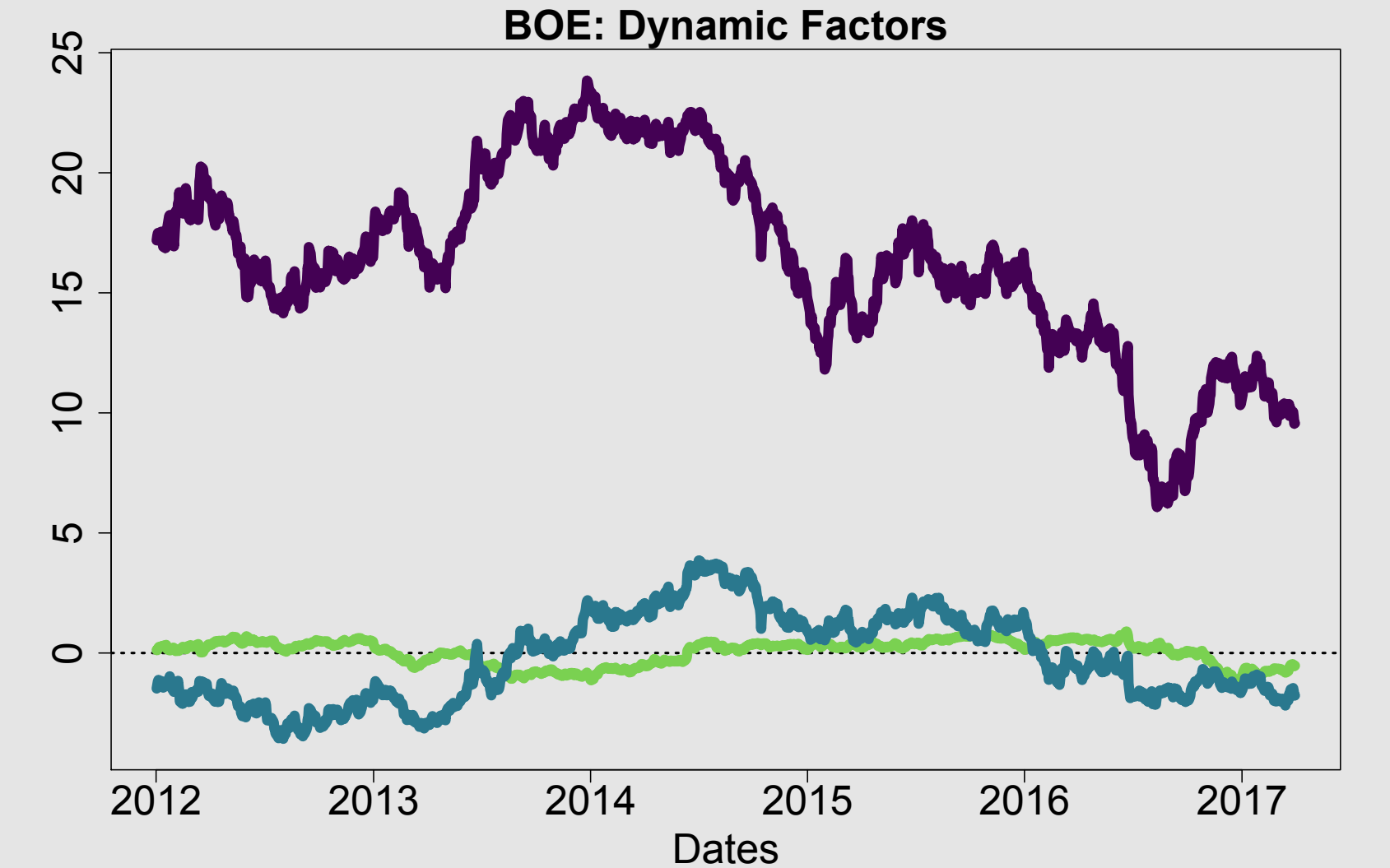
FITTED YIELD CURVE: FED



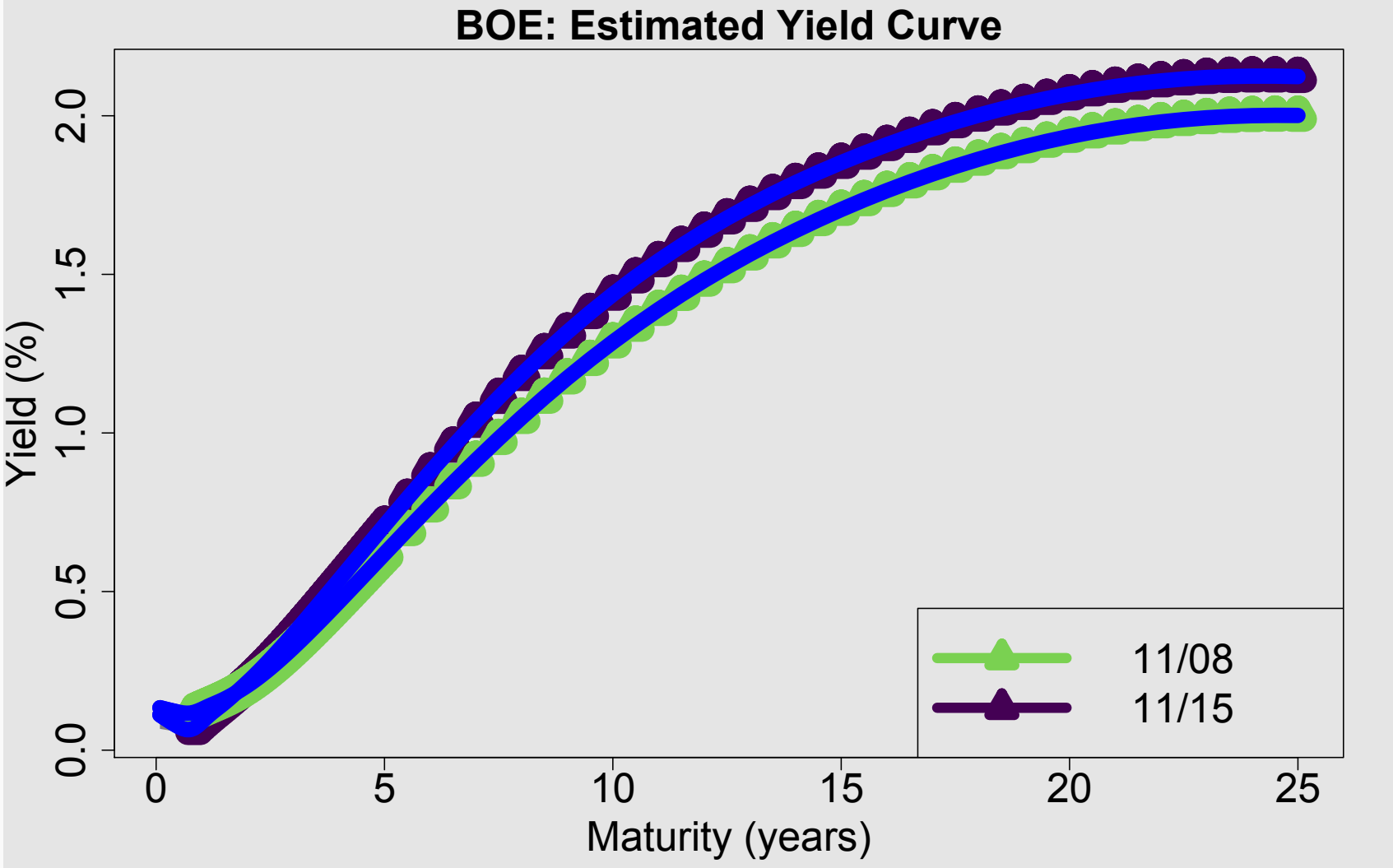
FACTOR LOADING CURVES: BOE



DYNAMIC FACTORS: BOE



FITTED YIELD CURVE: BOE



SUMMARY: CONTRIBUTIONS OF THE MFDLM

1. Hierarchical framework for modeling a multivariate time series of functional data
    - Covariates, stochastic volatility, change points
2. Joint estimation of model parameters; “exact” inference
3. Efficient Gibbs sampling algorithm; package `FDLM`
4. Applications: finance/economics and neuroscience; astronomy

REFERENCES

Helske, J. (2017).
KFAS: Kalman Filter and Smoother for Exponential Family State Space Models.
R package version 1.2.6.

Kowal, D. R., Matteson, D. S., and Ruppert, D. (2016).
A Bayesian multivariate functional dynamic linear model.
Journal of the American Statistical Association.
(in press).