University of Hawaii Economic Research Organization &  
Department of Economics  
University of Hawai‘i at Manoa

Macroeconomic Forecasting  
in the Era of Big Data

Peter Fuleky

AUBER  
Salt Lake City, Utah  
October 14, 2018
Let $X$ be a $mT \times K$ matrix containing observations on $K$ variables over $T$ years sampled at frequency $m$. Then

**Tall data:** $T \to \infty$ (long calendar span of data).

**Wide data:** $K \to \infty$ (large number of regressors).

**Dense data:** $m \to \infty$ (high-frequency intra-year sampling, regardless of whether the data are tall).
Dynamic Factor Models
Small and Large

Exact (small) DFMs:

\[ y_t = M f_t + \epsilon_t \]
\[ f_t = T f_{t-1} + \eta_t \]
\[ E(\epsilon_t \epsilon_{jt}) = 0 \quad \text{and} \quad E(\eta_t \eta_{jt}) = 0, \text{for} \ i \neq j \]

Approximate (large) DFMs:

\[ y_t = \beta' F_t + \gamma(L) y_{t-1} + \epsilon_t \]
\[ X_t = \Lambda F_t + \eta_t \]
\[ \lim_{n \to 1} \sum_{i=1}^{n} \sum_{j=1}^{n} |E(\eta_i \eta_{jt})| < \infty \]
Uncorrelated $F_t$ and $\eta_t$ imply:

$$\Sigma_{XX} = \Lambda \Sigma_{FF} \Lambda' + \Sigma_{\eta \eta}$$

Estimate the factors and loadings via principal components (NLLS):

$$\min_{F_1, \ldots, F_T, \Lambda} \frac{1}{T} \sum_{t=1}^{T} (X_t - \Lambda F_t)'(X_t - \Lambda F_t) \quad \text{s.t.} \quad \Lambda' \Lambda = I$$

If $\Sigma_{\eta \eta}$ is non-diagonal and $\eta$ is serially correlated use whitened $X$ and weighted (generalized) principal components (GLS).

Choose the number of factors based on information criteria.
1. Estimate the factors, $F_t$, by principal components.
2. Estimate the factor loadings, $\Lambda$, from a regression of $X_t$ on $\hat{F}_t$.
3. Estimate an autoregressive model for the residuals, $\hat{\eta}$, to obtain coefficients for idiosyncratic dynamics.
4. Estimate a vector autoregression of $\hat{F}_t$ on its lags, to obtain coefficients in the transition matrix.
5. Estimate the variance of the residuals of the vector autoregression.
6. Use the estimated coefficients in the Kalman smoother to obtain smoothed factors, $\tilde{F}_t$.
7. Project, $y_t$ on $\tilde{F}_{t-h}$ to estimate $\beta$ and make a direct forecast $h$ steps ahead.
DFM  The Kalman filter natively handles missing observations.

MIDAS  Distributed lag models can be used to model relationships between low frequency and high frequency variables.

\[
y_t^{(q)} = \beta_0 + \beta_1 b(L_m; \theta)x_{t+w-h}^{(m)} + \epsilon_t
\]

or UMIDAS:

\[
y_t^{(q)} = \delta_1(L_m)x_{t+w-h}^{(m)} + \epsilon_t
\]
Targeted Predictors

Before extracting factors from the data, narrow down $X$ using:

**Hard threshold** retain $X_i$ if its $t$ statistics in individual regression of $y$ on $X_i$ exceeds threshold at significance level $\alpha$.

**Soft threshold** based on penalized regressions:

\[
\text{Ridge regression: } \min_{\beta} RSS + \lambda \sum_{j=1}^{N} \beta_j^2
\]

Least absolute shrinkage selection operator: \[
\min_{\beta} RSS + \lambda \sum_{j=1}^{N} |\beta_j|
\]

Elastic net: \[
\min_{\beta} RSS + \lambda_1 \sum_{j=1}^{N} |\beta_j| + \lambda_2 \sum_{j=1}^{N} \beta_j^2
\]

Least angle regression: similar to forward stepwise regression.
1. Extract the $r$ largest common factors from the $T \times N$ matrix $X$ of potential predictors.

2. Construct bootstrap samples $(y_{1+h}^*, \hat{f}_1^*) \ldots (y_T^*, \hat{f}_{T-h}^*)$ by drawing with replacement blocks of $m$ rows of the dataset.

3. Orthogonalize the bootstrap factor draws.

4. Estimate the loading coefficients, discard the insignificant factors, re-estimate the model, and predict $\hat{y}_{t+h}^*$.

5. The bagged forecast is the average of $B$ bootstrap replications.
To forecast $y_t$ consider the predictors $z_t = (Z_t, Z_{t-1} \ldots Z_{t-p_{max}})$, with $Z_t = (y_{t-1}, F_{t1} \ldots F_{tf}, F_{t1}^2 \ldots F_{tr}^2)$.

**Component-wise $L_2$ boost** selects one predictor at a time.

1. Set $\hat{\Phi}_{t,0} = \bar{y}$ for each $t$.
2. For $m = 1 \ldots M$:
   a. let $u_t = y_t - \hat{\Phi}_{t,m-1}$ be the current residuals;
   b. regress $u_t$ on each $z_i$, select one that minimizes SSR;
   c. let $\hat{\phi}_m = z_{i_m^*}\hat{b}_{i_m^*}$;
   d. update $\hat{\Phi}_{t,m} = \hat{\Phi}_{t,m-1} + \nu\hat{\phi}_{t,m}$, where $0 < \nu \leq 1$.
3. The in-sample fit is $\hat{\Phi}_m(z) = \bar{y} + z'\hat{\beta}_m$, with recursion $\hat{\beta}_m = \hat{\beta}_{m-1} + \nu\hat{b}_m^*$, where $\hat{b}_m^* \neq 0$ only in the $i^{th}$ position. At the final step, $\hat{\beta}_M$ will likely have many zero elements,

**Block-wise $L_2$ boost** selects the predictor and its lags jointly at each iteration.
Macroeconomic Forecasting in the Era of Big Data - Table of Contents

Link: http://www2.hawaii.edu/~fuleky/BigDataSite/index.html

Introduction  Big Data Sources and Types.

Capturing Relationships  Dynamic Factor Models; Factor Augmented Vector Autoregressions, Panel VARs, and Global VARs; Bayesian Vector Autoregressions; Mixed Frequency Data Sampling Regressions; Neural Nets.

Seeking Parsimony  Penalized Regression; Estimation of Common Factors; Subspace Methods; Variable Selection and Feature Screening; Robust Variable Selection, Regression, and Covariance Estimation.

Dealing with Model Uncertainty  Frequentist Averaging; Bayesian Averaging; Bootstrap Aggregation; Cross-validation Aggregation; Boosted Regression Trees.

Further Issues  Unit Roots and Cointegration; Time Varying Parameters; Turning Points and Classification; Volatility Forecasts; Density Forecasts; Frequency Domain; Hierarchical Time Series; Forecast Evaluation.
References I


