Discussion
Macroeconomic forecasting in times of crisis
Pablo Guerron-Quintana and Molin Zhong

Srečko Zimic*

*European Central Bank

September, 2017
(Time-series) Forecasting = use historical patterns to predict future
(Time-series) Forecasting = use historical patterns to predict future

- Linear models consider (average) pattern.
(Time-series) Forecasting = use historical patterns to predict future

- Linear models consider (average) pattern.
- Times of crisis ⇒ different times!
- Only around 10 percent of the time economy is in recession
(Time-series) Forecasting = use historical patterns to predict future

- Linear models consider (average) pattern.
- Times of crisis ⇒ different times!
- Only around 10 percent of the time economy is in recession

Solution: use (crisis) patterns to forecast during crisis?
How to define and find the patterns and how to use them?

Match the current time series with the "most equal" pattern in history.

- Cut the data into blocks of length $k$.
- Compare the current block with all blocks via distance function:

$$
dist = \sum_{i=1}^{k} w(i)(y_{T-k+i} - y_i)^2
$$

- Only the closest blocks provide information for the forecast.

Assume the match:

- Current data block: $B^C = y_T - k, \ldots, y_T, y_T$
- Best match data block: $B^1 = y_1, \ldots, y_{k-1}, y_k$

To forecast $y_{T+1}$ we use information contained in $y_{k+1}$ (and $B^1$).
Framework

- Completely non-parametric approach: \( \hat{y}_{T+1} = y_{k+1} \)
- In the paper semi-parametric approach:

\[
\hat{y}_{T+1} = (y_{k+1} - \hat{y}_{k+1,ARIMA}) + \hat{y}_{T+1,ARIMA}
\]

- \( \hat{y}_{T+1,ARIMA} \): parametric ARIMA forecast.
- \( (y_{k+1} - \hat{y}_{k+1,ARIMA}) \): correction for forecast error made by ARMA model in “similar” period.

- Match with \( m \) similar periods:

\[
\hat{y}_{T+1} = \frac{1}{m} \sum_{i}^{m} (y_{l(i)+1} - \hat{y}_{l(i)+1,ARIMA}) + \hat{y}_{T+1,ARIMA}
\]

- Machine learning step: estimate two parameters, \( k \) and \( m \).
- Select \( k \) and \( m \) by minimizing out-of-sample forecast error.
Other variables may provide “pattern” information.

Reinhart and Rogoff (2014): Financial crisis $\Rightarrow$ protracted and halting nature of the recovery

Compare multivariate block, including financial variables:

$$\text{dist} = \sum_{i=1}^{k} w(i)((x_{T-k+i} - x_i)^2 + (z_{T-k+i} - z_i)^2)$$

Financial variables provide important information to identify patterns.
What is the computational burden to estimate two parameters?

In principle one could estimate more parameters:

- The weights for additional series.
- The weights for different blocks.
- The weighting function for lags.
- The weight on parametric vs. non-parametric forecast:

\[ \hat{y}_{T+1} = w y_{k+1} + (1 - w) \hat{y}_{T+1, ARIMA} \]
Comments: Real-time vs. revised series

- Forecasting evaluation done with last vintage (revised data).
- Real-Time estimate of Industrial production (SPF) growth vs. revised estimate:

In Real-Time harder to capture changing patterns!

Leading variables (financial variables) could be potentially even more useful.
Can this methodology be used also to produce density forecasts? Given that multiple blocks are matched this seems natural:

\[
\hat{y}_{T+1,i} = (y_{l(i)+1} - \hat{y}_{l(i)+1, ARIMA}) + \hat{y}_{T+1, ARIMA}
\]