Financial Econometrics Econ 40357
Introduction to Forecasting

N.C. Mark

University of Notre Dame and NBER

September 7, 2020

Forecast evaluation

Time map

$$\underbrace{\underbrace{1,2,\ldots,\overbrace{t}^{\text{today}}}_{\text{In sample data}}}_{\text{initial estimation}},\ \overbrace{\underbrace{\underbrace{t+1}_{\text{1-step ahead}},\ t+2,\ \underbrace{t+3}_{\text{3-steps ahead}}}_{\text{Multi-step ahead forecast}},\ t+4,\ldots,T}^{\text{Reserved for forecast evaluation}}$$

- In EViews, Static forecasts give a sequence of one-step ahead forecasts.
- Dynamic forecasts give the multi-period ahead forecasts.
- The estimated parameters are not automatically updated

Forecasts (conditional expectation) and forecast errors

- Illustrate ideas with ARMA(2,1) model (deviation from mean)

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \theta \epsilon_{t-1} + \epsilon_t$$

- One-step ahead forecast (conditional expectation)

$$\hat{y}_{t+1|t} = \hat{\rho}_1 y_t + \hat{\rho}_2 y_{t-1} + \hat{\theta} \hat{\epsilon}_t$$

- One-step ahead forecast error

$$fe_{t+1|t} = y_{t+1} - \hat{y}_{t+1|t}$$

- Two-step ahead forecast

$$\hat{y}_{t+2|t} = \hat{\rho}_1 \hat{y}_{t+1|t} + \hat{\rho}_2 y_t + \underbrace{\hat{\theta} \hat{\epsilon}_{t+1|t}}_{0} = \hat{\rho}_1 \hat{y}_{t+1|t} + \hat{\rho}_2 y_t$$

- Two-step ahead forecast error

$$fe_{t+2|t} = y_{t+2} - \hat{y}_{t+2|t}$$

- Similarly, for $k-$step ahead forecasts and forecast errors.

Random Walk Forecast

Random walk with drift

$$\begin{aligned} y_{t+1} &= \mu + y_t + \epsilon_t \\ \hat{y}_{t+1|t} &= \hat{\mu} + y_t \\ \hat{y}_{t+2|t} &= \hat{\mu} + \hat{y}_{t+1|t} = 2\hat{\mu} + y_t \\ &\vdots \\ \hat{y}_{t+k|t} &= k\hat{\mu} + y_t \end{aligned}$$

Driftless Random Walk (no change) Forecast

Driftless random walk, set $\mu = 0$.

$$\begin{aligned}
y_{t+1} &= y_t + \epsilon_t \\
\hat{y}_{t+1|t} &= y_t \\
\hat{y}_{t+2|t} &= \hat{y}_{t+1|t} = y_t \\
&\vdots \\
\hat{y}_{t+k|t} &= y_t
\end{aligned}$$

The no-change forecast.

$$\hat{y}_{t+1|t} - y_t = 0$$

This model requires no estimation. Can your model beat the no-change forecast?

Estimation windows

- Estimate over fixed sub-sample. $t = 1, \ldots, t_0$. Forecst $\hat{y}_{t_0+1|t_0}, \hat{y}_{t_0+2|t_{0+1}}, \ldots \hat{y}_{T|T-1}$
- Estimate over full sample. $t = 1, \ldots, T$. This produces 'in-sample' forecasts. These are pseudo forecasts because they use out-of-sample information. Cheating
- Recursive updated estimates–'Out-of-sample' forecasts.
  - Estimate from $t = 1, \ldots, t_0$, forecast $\hat{y}_{t_0+1|t_0}$.
  - Re-estimate from $t = 1, \ldots, t_0 + 1$, forecast $\hat{y}_{t_0+2|t_0+1}$, and so on until,
  - Estimate from $t = 1, ..., T - 1$, forecast $\hat{y}_{T|T-1}$.
- Rolling sample estimates.
  - Estimate $t = 1, ..., t_0$, forecast $\hat{y}_{t_0+1|t_0}$.
  - Estimate $t = 2, ..., t_0 + 1$, forecast $\hat{y}_{t_0+2|t_0+1}$, and so on until
  - Estimate $t = T - t_0 + 1, ..., T - 1$, forecast $\hat{y}_{T|T-1}$ Do this if you suspect substantial parameter instability

Evaluate forecast accuracy

- You have $T - t_1 + 1$ forecast errors, generated from your preferred model and the random walk item.
- Compute accuracy measures for both models and compare.
- Root-mean-square forecast errors (RMSFE).

$$\text{RMSFE} = \sqrt{\frac{1}{T - t_1 + 1} \sum_{t=t_1}^{T} fe_{t,t-1}^2} \tag{1}$$

- Mean absolute forecast errors (MAFE).

$$\text{MAFE} = \frac{1}{T - t_1 + 1} \sum_{t=t_1}^{T} \left| fe_{t,t-1} \right| \tag{2}$$

Evaluate forecast accuracy

Theil's U statistic: Compare your forecasts to the 'no-change' forecast.

$$U = \frac{RMSFE_{\text{Model}}}{RMSFE_{\text{NC}}} \tag{3}$$

If $U < 1$, you are beating the no-change forecast. Your model has some predictive content.

Evaluate forecast accuracy

Money making ability (economic significance) Sign prediction (buy-sell signals). For returns, let z_t be such that

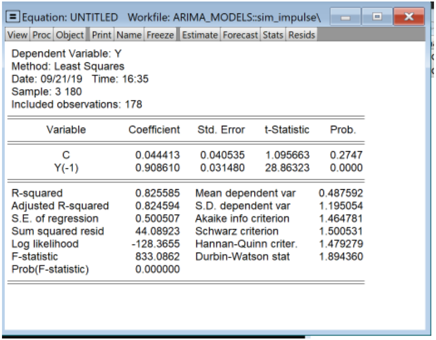
$$z_t = \begin{cases} 1 & \text{if } (y_{t+1}\hat{y}_{t+1}) > 0 \text{ (same sign)} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

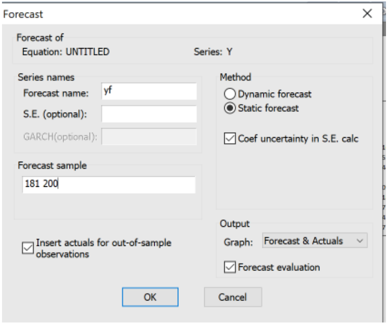
$$\begin{array}{c}\text{Percentage forecasts} \\ \text{with same sign}\end{array} = \frac{100}{T - t_1} \sum_{t=t_1}^{T-1} z_t$$

Your model will be ‘successful’ if it exceeds 50%

Eviews example

200 observations. Estimate AR(1) on observations 3 to 180. Ask for static forecasts of observations 181 to 200





Eviews example