3. Forecasting Basics

author: ECONOMIC FORECASTING date: Summer 2017 autosize: true

==

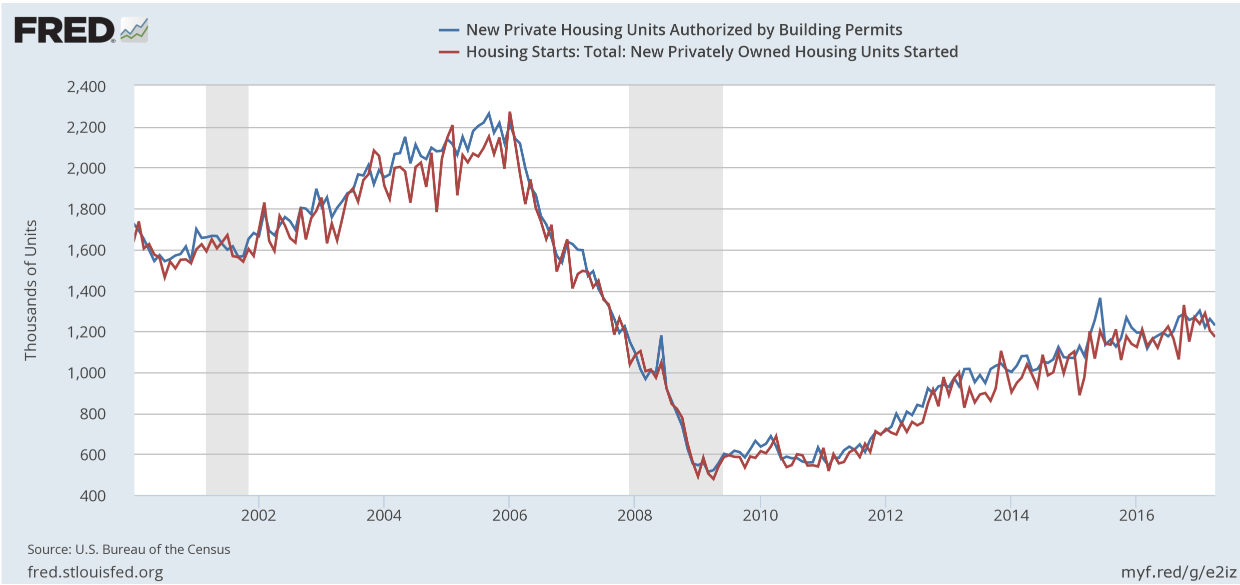

Figure 1:

Leading Variables

Examples for GDP: - Average weekly number of initial applications for unemployment insurance - Consumer sentiment - S&P500 Stock Index - Number of new orders of capital goods unrelated to defense - Inflation-adjusted money supply (M2) - Speed of delivery of new merchandise to vendors from suppliers

Cross Correlation

Suppose x and y have constant expected values and variances. The **cross covariance** function $(\gamma_k(x,y))$ is:

$$\gamma_k(x,y) = \mathbf{E}[(x_{t+k} - \mu_x)(y_t - \mu_y)]$$

and the **cross correlation** function is:

$$\rho_k(x,y) = \frac{\gamma_k(x,y)}{\sigma_x \sigma_y}$$

Sample Cross Correlation

And the **sample cross correlation** function is:

$$c_k(x,y) = \frac{1}{n}\sum_{t=1}^{n-k}(x_{t+k} - \bar{x})(y_t - \bar{y})$$

and the **sample acf and cross correlation** is:

$$r_k(x,y) = \frac{c_k(x,y)}{\sqrt{c_0(x,x)c_0(y,y)}}$$

Cross Correlogram

In R, `acf(ts.union())`

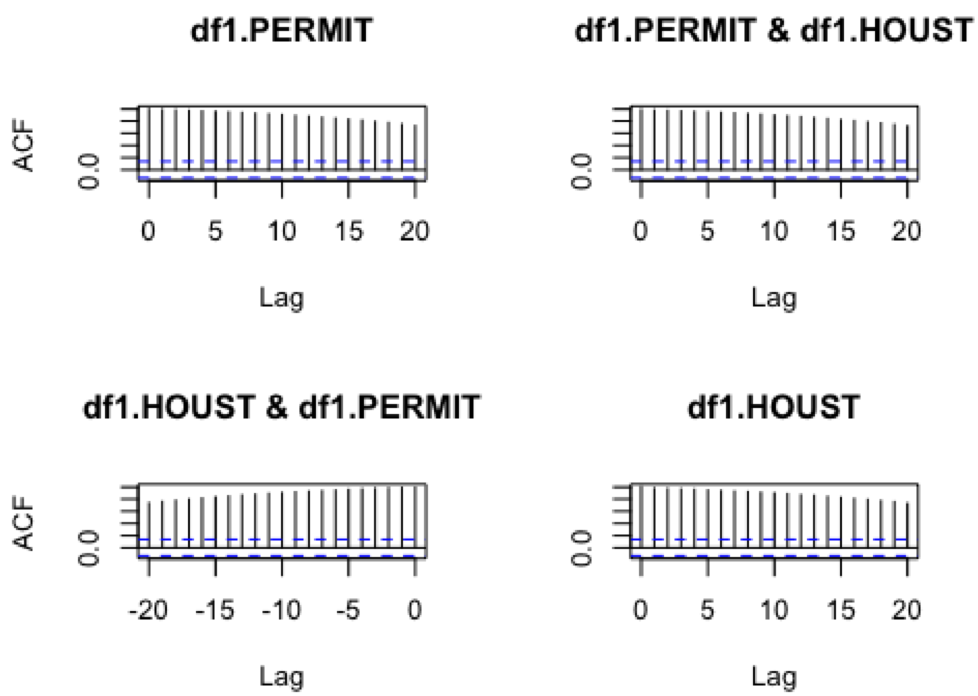

Figure 2:

Choosing a Forecasting Technique

Factors to Consider - Horizon - Turning Points - Deadlines - Ease of understanding/explaining - Results of an Empirical Evaluation

Empirical Evaluation

More complex models may be better at predicting the past, but not necessarily better at predicting the future.

Forecast Error: - Compare accuracy between techniques - Measure reliability - Help search for optimal technique - Residual: $y - \hat{y} = e$ + difference between forecast and actual value

Measuring Forecast Error

Mean Absolute Deviation (MAD) - Average of magnitude of forecasting errors - Same units as original data

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

Mean Squared Error (MSE) - Penalizes large forecasting error

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

Measuring Forecast Error

Root Mean Squared Error (RMSE) - Same units as original data - Penalizes large forecasting error

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

Mean Absolute Percentage Error (MAPE) - Converts error to a percentage (useful when y_t is large)

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t}\right) \cdot 100$$

Loss Functions

An additional factor to consider when evaluating a model/forecast is the loss incurred due to decisions made from the model/forecast

Example:

A company has a choice - to replenish or deplete their inventory - if demand is high, they would like to replenish - if demand is low, they would rather deplete

==

Symmetric Loss

	High Demand	Low Demand
Replenish	\$0	\$10,000
Deplete	\$10,000	\$0

Asymmetric Loss

	High Demand	Low Demand
Replenish	\$0	\$10,000
Deplete	\$20,000	\$0

Loss Function

A **Loss Function** $\mathbf{L}(e)$ where $e_t = y_t - \hat{y}_t$, must satisfy 3 conditions.

1. $\mathbf{L}(0) = 0 \rightarrow$ no loss when the forecast error is 0.
2. $\mathbf{L}(e)$ is continuous
3. $\mathbf{L}(e)$ is increasing on both sides of the origin. The bigger the error, the bigger the loss.

Aside from the above restrictions, a loss function can take on any form.

Direction of Change Loss

$$\mathbf{L}(y, \hat{y}) = \begin{cases} 0, & \text{if sign } (\delta y) = \text{ sign } (\delta \hat{y}) \\ 1, & \text{if sign } (\delta y) \neq \text{ sign } (\delta \hat{y}) \end{cases}$$

With the above loss function, if you predict the direction of the change correctly, you incur no loss; but if your prediction is wrong, you're penalized.

Naive Method

Naive Model:

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

Assumes most recent period is the best predictor of the future. - Does NOT consider the past

What kind of data would this method be used for?

Naive Continued...

What if the data have a trend? - If data is trending upwards, the naive model will consistently under-predict (downward bias)

Enter the **Naive Trend Model**:

$$\begin{aligned} \hat{y}_{t+1} &= y_t + \delta y_t \\ &= y_t + (y_t - y_{t-1}) \\ &= 2y_t - y_{t-1} \end{aligned}$$

Seasonal Data

What if the data are seasonal? - Depending on which part is to be forecast, errors could be large or small

Enter the **Naive Seasonal Model** (Quarterly):

$$\hat{y}_{t+1} = y_{t-3}$$

With trend:

$$\hat{y}_{t+1} = y_{t-3} + \frac{y_t - y_{t-4}}{4}$$

Simple Techniques Wrap-Up

PROS - Simple to use - Quick to compute - Adapts quickly to new data - Can be used as a baseline to compare more sophisticated techniques

CONS - Ignores historical data - Random fluctuations and fundamental changes have equal weight.

Averaging Methods

Instead of using all observations, a forecaster can use an averaging method that uses only the more recent observations in its forecast.

Moving Average (MA)

$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + ... + y_{t-k+1}}{k}$$

where k is the number of terms in the moving average. - MA(1) is equivalent to the Naive Model - MA(t) is the simple average

Double Moveing Average

-The average of the moving average

Double Moving Average (M)

$$M_t' = \frac{M_t + M_{t-1} + ... + M_{t-k+1}}{k}$$

where

$$M_t = \hat{y}_{t+1} = \frac{y_t + y_{t-1} + ... + y_{t-k+1}}{k}$$

helps forecast trend data

Moving Average with Trend

Combining the moving average and double moving average to produce a forecast

"Intercept"

$$a_t = M_t + (M_t - M_t') = 2M_t - M_t'$$

"Slope"

$$b_t = \frac{2}{k-1}(M_t - M_t')$$

Forecast

$$\hat{y}_{t+p} = a_t + b_t p$$

where p is the number of periods to be forecast.

Averaging Methods Wrap-Up

PROS - Fairly easy to use - Can be used to predict stationary or trending data

CONS - Depending on the number of terms in the moving average, the moving average can be slow to react to structural changes - Judgment must be used to predict the number of terms in the moving average

Exponential Smoothing

Exponential smoothing uses all the observations in producing the forecast - However, recent observations are weighted more heavily than old observations - Forecast is continually updated as more observations are considered

Goal: Estimate the current level and then use this estimate as the forecast for future values

Exponential Smoothing

Exponential Smoothing:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t \ \ (0 \leq \alpha \leq 1)$$

and expanding this…

$$\hat{y}_{t+1} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \cdots$$

α determines the speed at which old observations lose their impact on the forecasted value.

What happens as $\alpha \to 1$?

Applying the Smoothing Equation

Recall...

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t$$

Where does the initial level estimate come from? - Typically, the initial level estimate is simply the average of the first half of the data - This becomes the smoothed estimate at $t1$, which is the estimate for $t2$.

Exponential Trend with Smoothing

Rarely do data series exhibit constant trending behavior

Linear exponential smoothing (Holt) allows us to model trending data that changes over time

Goal: Estimate current level *and* slope, then use these estimates to allow us to forecast future values

Smoothing with Trend Model

Current Level Estimate

$$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$$

Current Trend Estimate

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

Forecast

$$\hat{y}_{t+p} = L_t + pT_t$$

Smoothing with Seasonality

Winter's method is a multiplicative smoothing method that allows us to use a smoothing method that has the ability to control for trend and seasonality

Goal: Estimate current level, trend, *and* seasonality and use these estiamates to forecast future values.

Smoothing with Seasonality Model

Current Level Estimate

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

Current Trend Estimate

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

Current Seasonality Estimate

$$S_t = \gamma \frac{y_t}{L_t} + (1-\gamma)S_{t-s}$$

Forecast

$$\hat{y}_{t+p} = (L_t + pT_t)S_{t-s+p}$$

In R

Use `HoltWinters()`

- Level only: set `beta=FALSE` and `gamma=FALSE`
- Adding trend: set `gamma=FALSE`
- Full model: allow `alpha`, `beta`, and `gamma` to be estimated
  - additionally, you can choose between additive and multiplicative seasonality

For predictions, use `predict()` with the `HoltWinters` outcome as the as the `object` in `predict()`

Exponential Smoothing Wrap-Up

PROS - Fairly quick, low-cost estimation - When done properly, can provide good short-run forecasts

CONS - Atheoretical - Choices of $\alpha, \beta,$ and γ