Forecasting Economic Time Series Using Statgraphics Centurion

Presented by Dr. Neil W. Polhemus
Time Series

• “A sequence of numerical data points in successive order, using occurring in uniform intervals.” – www.investopedia.com

• Examples
  – Daily closing stock prices
  – Monthly unemployment rates
  – Quarterly GDP

• Notation: \{Y_t\}, \ t = 1, 2, \ldots, n
Example – U.S. Quarterly GDP
Time Series Components

- Trend
- Cycle
- Seasonality
- Random or irregular component
Trend Analysis

$GDP^{0.5} = 40.3495 + 0.339444 \times \text{Quarter}$

RMSE: 324.724, $R^2$: 99.30%, P-Value: 0.0000
Differencing Operators

- First Differences

\[ \nabla Y_t = Y_t - Y_{t-1} \]

- Second Differences

\[ \nabla^2 Y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \]
First Differences: \( \nabla Y_t = Y_t - Y_{t-1} \)
First Differences after Square Root
Types of Forecasting Models

- **Autoprojective models** – models that involve only the time series to be forecast. These models capture the dynamics of past time series movements and project them into the future.

- **Models with leading indicators** – models that include past values of other time series variables.
GDP and New Construction Permits
Notation

• Time series to be forecast:

\[ \{Y_t\}, \ t = 1, 2, 3, \ldots, n \]

• Forecasts:

\[ F_t(k) = \text{forecast of } Y_{t+k} \text{ using information available at time } t \]

• One-ahead forecast errors:

\[ \hat{\varepsilon}_t = Y_t - F_{t-1}(1) \]
Types of Autoprojective Models

1. **Random walk** - current value has all relevant information.

   without constant: \( F_t(k) = Y_t \) for all \( k \geq 1 \)

   with constant: \( F_t(k) = Y_t + k\hat{\Delta} \)

   where \( \hat{\Delta} \) is mean difference between consecutive periods
Types of Autoprojective Models

2. **Trend models** – time series follows a deterministic trend with random fluctuations around the trend.

\[ F_t(k) = \hat{a} + \hat{b}(t + k) \]

\[ F_t(k) = \exp\left(\hat{a} + \hat{b}(t + k)\right) \]

\[ F_t(k) = \exp\left(\hat{a} + \hat{b} / (t + k)\right) \]
Types of Autoprojective Models


\[ F_t(k) = \frac{\sum_{i=0}^{c-1} Y_{t-i}}{c} \]
Types of Autoprojective Models

4. **Exponential smoothing** – combines new information with previous forecasts to generate new forecasts.

\[ F_t(k) = \alpha Y_t + (1 - \alpha) F_{t-1}(1) \]

Statgraphics has simple, linear, quadratic and seasonal smoothers.
Holt’s Linear Exp. Smoothing
Types of Autoprojective Models

5. **ARIMA Models** – parametric models which describe system dynamics.

ARIMA($p,d,q$) model has:

- autoregressive term of order $p$
- moving average term of order $q$
- applied to the differences of order $d$
Autoregressive Models

- **AR(1)**

\[ Y_t = \mu + \phi_1(Y_{t-1} - \mu) + \varepsilon_t \]

- **AR(2)**

\[ Y_t = \mu + \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \varepsilon_t \]
Moving Average Models

• MA(1)

\[ Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \]

• MA(2)

\[ Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \]
ARMA Models

- ARMA(1,1)

\[ Y_t = \mu + \phi_1 (Y_{t-1} - \mu) + \epsilon_t - \theta_1 \epsilon_{t-1} \]
ARIMA Models

- ARIMA(1,1,1)

\[ \nabla Y_t = Y_t - Y_{t-1} \]

\[ \nabla Y_t = \mu + \phi_1 (\nabla Y_{t-1} - \mu) + \varepsilon_t - \theta_1 \varepsilon_{t-1} \]

Note: \( \mu \) is sometimes omitted.
Automatic Forecasting

Automatic Forecasting

Data:
- GDP

(Time Indices):
- Quarter

(Sampling Interval):
- Once Every: 1
- Year(s) [4-digit]
- Quarter(s)
- Month(s)
- Day(s)

(Seasonality):

(Trading Days Adjustment):

(Number of Forecasts: 3
Withhold for Validation: 0

Sort column names

OK  Cancel  Delete  Transform  Help

statgraphics® centurion
Analysis Options
Method Selection Criterion

Akaike Information Criterion

\[
AIC = 2 \ln(RMSE) + \frac{2c}{n}
\]

c = number of coefficients in fitted model

RMSE = root mean squared error calculated from the one-period ahead forecast errors
Adjustments
Adjustments

1. **Trading days adjustment** – used to normalize monthly data by dividing each data value by number of trading days in the month.

2. **Math adjustment** – transforms each data value before fitting models.

3. **Seasonal adjustment** – removes seasonal effects using seasonal decomposition prior to fitting models.

4. **Inflation adjustment** – corrects each data value for a constant rate of inflation.
Analysis Summary

**Automatic Forecasting - GDP**

Data variable: GDP (billions of chained 2009 dollars, seasonally adjusted)

Number of observations = 221
Time indices: Quarter (from BEA)

**Forecast Summary**

- Math adjustment: Square root
- Forecast model selected: ARIMA(2,1,0) with constant
- Number of forecasts generated: 3
- Number of periods withheld for validation: 0

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Estimation</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>64.215</td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>46.0611</td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>0.572054</td>
<td></td>
</tr>
<tr>
<td>ME</td>
<td>-0.00425189</td>
<td></td>
</tr>
<tr>
<td>MPE</td>
<td>-0.00142829</td>
<td></td>
</tr>
</tbody>
</table>

**ARIMA Model Summary**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.28558</td>
<td>0.0672908</td>
<td>4.24396</td>
<td>0.0000033</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.145362</td>
<td>0.0672999</td>
<td>2.15991</td>
<td>0.031876</td>
</tr>
<tr>
<td>Mean</td>
<td>0.32359</td>
<td>0.0391041</td>
<td>8.27509</td>
<td>0.000000</td>
</tr>
<tr>
<td>Constant</td>
<td>0.184141</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Backforecasting: yes

Estimated white noise variance = 0.111779 with 217 degrees of freedom
Estimated white noise standard deviation = 0.334334
Number of iterations: 1
Model Comparisons

Model Comparison
Data variable: GDP
Number of observations = 221

Models
(A) Random walk
(B) Random walk with drift = 0.32566
(C) Constant mean = 91.6055
(D) Linear trend = 53.9272 + 0.339444 t
(E) Quadratic trend = 55.3126 + 0.302169 t + 0.000167904 t^2
(F) Exponential trend = exp(4.06538 + 0.00380989 t)
(G) S-curve trend = exp(4.52201 + -1.24712 /t)
(H) Simple moving average of 2 terms
(I) Simple exponential smoothing with alpha = 0.9999
(J) Brown's linear exp. smoothing with alpha = 0.6898
(K) Holt's linear exp. smoothing with alpha = 0.9999 and beta = 0.0046
(L) Brown's quadratic exp. smoothing with alpha = 0.4852
(M) ARIMA(2,1,0) with constant
(N) ARIMA(1,1,1) with constant
(O) ARIMA(1,1,0) with constant
(P) ARIMA(0,1,2) with constant
(Q) ARIMA(2,1,1) with constant
### Model Statistics

<table>
<thead>
<tr>
<th>Estimation Period</th>
<th>Model</th>
<th>RMSE</th>
<th>MAE</th>
<th>MAPE</th>
<th>ME</th>
<th>MPE</th>
<th>AIC</th>
<th>HQC</th>
<th>SBIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>91.8364</td>
<td>76.2823</td>
<td>0.93793</td>
<td>59.7314</td>
<td>0.743744</td>
<td>9.04002</td>
<td>9.04002</td>
<td>9.04002</td>
<td></td>
</tr>
<tr>
<td>(B)</td>
<td>68.7816</td>
<td>48.0714</td>
<td>0.601831</td>
<td>0.0671557</td>
<td>-0.00743105</td>
<td>8.47092</td>
<td>8.47713</td>
<td>8.4863</td>
<td></td>
</tr>
<tr>
<td>(C)</td>
<td>4076.34</td>
<td>3501.1</td>
<td>48.6727</td>
<td>472.255</td>
<td>-19.5762</td>
<td>16.635</td>
<td>16.6412</td>
<td>16.6503</td>
<td></td>
</tr>
<tr>
<td>(D)</td>
<td>367.469</td>
<td>296.302</td>
<td>3.37438</td>
<td>3.30028</td>
<td>-0.0366323</td>
<td>11.8314</td>
<td>11.8438</td>
<td>11.8621</td>
<td></td>
</tr>
<tr>
<td>(E)</td>
<td>369.041</td>
<td>281.256</td>
<td>2.93507</td>
<td>2.92671</td>
<td>-0.0809698</td>
<td>11.849</td>
<td>11.8676</td>
<td>11.8951</td>
<td></td>
</tr>
<tr>
<td>(F)</td>
<td>589.219</td>
<td>391.496</td>
<td>4.05372</td>
<td>-20.4197</td>
<td>-0.127704</td>
<td>12.7757</td>
<td>12.7881</td>
<td>12.8064</td>
<td></td>
</tr>
<tr>
<td>(G)</td>
<td>3825.45</td>
<td>3207.88</td>
<td>40.476</td>
<td>846.503</td>
<td>-9.99178</td>
<td>16.517</td>
<td>16.5294</td>
<td>16.5477</td>
<td></td>
</tr>
<tr>
<td>(H)</td>
<td>126.82</td>
<td>106.906</td>
<td>1.30998</td>
<td>90.1888</td>
<td>1.11967</td>
<td>9.69458</td>
<td>9.70079</td>
<td>9.70996</td>
<td></td>
</tr>
<tr>
<td>(I)</td>
<td>91.8422</td>
<td>75.943</td>
<td>0.933758</td>
<td>59.4671</td>
<td>0.740453</td>
<td>9.04919</td>
<td>9.0554</td>
<td>9.06457</td>
<td></td>
</tr>
<tr>
<td>(J)</td>
<td>68.679</td>
<td>50.2451</td>
<td>0.624134</td>
<td>-0.290729</td>
<td>0.0072373</td>
<td>8.46794</td>
<td>8.4715</td>
<td>8.48331</td>
<td></td>
</tr>
<tr>
<td>(K)</td>
<td>71.1219</td>
<td>51.7957</td>
<td>0.651243</td>
<td>17.2005</td>
<td>0.246029</td>
<td>8.54689</td>
<td>8.55931</td>
<td>8.57764</td>
<td></td>
</tr>
<tr>
<td>(L)</td>
<td>76.6172</td>
<td>56.2215</td>
<td>0.702866</td>
<td>-1.16001</td>
<td>-0.0153938</td>
<td>8.68669</td>
<td>8.6929</td>
<td>8.70207</td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>64.215</td>
<td>46.0611</td>
<td>0.572054</td>
<td>-0.00425189</td>
<td>-0.00142829</td>
<td>8.35162</td>
<td>8.37025</td>
<td>8.39775</td>
<td></td>
</tr>
<tr>
<td>(N)</td>
<td>64.2656</td>
<td>46.193</td>
<td>0.574924</td>
<td>0.125512</td>
<td>0.00328955</td>
<td>8.3532</td>
<td>8.3782</td>
<td>8.39933</td>
<td></td>
</tr>
<tr>
<td>(O)</td>
<td>64.6991</td>
<td>46.524</td>
<td>0.577422</td>
<td>-0.0032121</td>
<td>-0.00330691</td>
<td>8.3576</td>
<td>8.37001</td>
<td>8.38835</td>
<td></td>
</tr>
<tr>
<td>(P)</td>
<td>64.587</td>
<td>46.1862</td>
<td>0.572778</td>
<td>0.0258493</td>
<td>-0.00296255</td>
<td>8.36318</td>
<td>8.3818</td>
<td>8.40931</td>
<td></td>
</tr>
<tr>
<td>(Q)</td>
<td>64.3437</td>
<td>46.0917</td>
<td>0.571796</td>
<td>0.0936972</td>
<td>-0.000419893</td>
<td>8.36468</td>
<td>8.38951</td>
<td>8.42618</td>
<td></td>
</tr>
</tbody>
</table>
# Model Residual Analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>RUNS</th>
<th>RUNM</th>
<th>AUTO</th>
<th>MEAN</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>91.8364</td>
<td>OK</td>
<td>**</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(B)</td>
<td>68.7816</td>
<td>OK</td>
<td>**</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(C)</td>
<td>4076.34</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>OK</td>
</tr>
<tr>
<td>(D)</td>
<td>367.469</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(E)</td>
<td>369.041</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>*</td>
<td>***</td>
</tr>
<tr>
<td>(F)</td>
<td>589.219</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>OK</td>
<td>***</td>
</tr>
<tr>
<td>(G)</td>
<td>3825.45</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>(H)</td>
<td>126.82</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(I)</td>
<td>91.8422</td>
<td>OK</td>
<td>**</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(J)</td>
<td>68.679</td>
<td>OK</td>
<td>OK</td>
<td>*</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(K)</td>
<td>71.1219</td>
<td>OK</td>
<td>OK</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(L)</td>
<td>76.6172</td>
<td>OK</td>
<td>OK</td>
<td>***</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(M)</td>
<td>64.215</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(N)</td>
<td>64.2656</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(O)</td>
<td>64.6991</td>
<td>*</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(P)</td>
<td>64.587</td>
<td>*</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>(Q)</td>
<td>64.3437</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
</tbody>
</table>

**Key:**

- RMSE = Root Mean Squared Error
- RUNS = Test for excessive runs up and down
- RUNM = Test for excessive runs above and below median
- AUTO = Box-Pierce test for excessive autocorrelation
- MEAN = Test for difference in mean 1st half to 2nd half
- VAR = Test for difference in variance 1st half to 2nd half
- OK = not significant (p >= 0.05)
- * = marginally significant (0.01 < p <= 0.05)
- ** = significant (0.001 < p <= 0.01)
- *** = highly significant (p <= 0.001)
Time Sequence Plot
Forecast Plot

Forecast Plot for GDP
ARIMA(2,1,0) with constant
- actual
- forecast
- 95.0% limits

GDP

Q1/14  | Q1/15  | Q1/16
16000  | 16200  | 16400  
16600  | 16800  | 17000  
Quarter
### Forecast Table

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower 95.0%</th>
<th>Upper 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2/15</td>
<td>16315.3</td>
<td>16147.4</td>
<td>16484.1</td>
</tr>
<tr>
<td>Q3/15</td>
<td>16372.6</td>
<td>16099.1</td>
<td>16648.4</td>
</tr>
<tr>
<td>Q4/15</td>
<td>16443.6</td>
<td>16070.2</td>
<td>16821.4</td>
</tr>
</tbody>
</table>
Residual ACF

Residual Autocorrelations for adjusted GDP
ARIMA(2,1,0) with constant

Autocorrelations vs. lag
Residual Crosscorrelations

Estimated Crosscorrelations for Residuals with DIFF(Permits)
ARIMA(2,1,0) with constant
Models with Leading Indicators

• The ARIMA model is modified by adding additional terms involving one or more regressors \{X_t\}.

• The same differencing and AR operators are applied to \{X_t\} as are applied to \{Y_t\}.

• We are essentially fitting an ARIMA model to the errors of the regression of Y on X.

• Helpful discussion of this by Prof. Robert Nau at people.duke.edu/~rnau/arimreg.htm
One Complication

To use a time series such as *Permits* in our forecast model:

– Generate a forecasting model for the regressor variable(s).
– Add the forecasts to the bottom of the datasheet.
– Add the regressors to our model using Analysis Options.
Forecasting Construction Permits

Automatic Forecasting

Data:
- Permits

(Time Indices):
- Quarter

Sampling Interval:
- Once Every: 1

Seasonality:

Trading Days Adjustment:

Select:

Sort column names

Number of Forecasts: 3
Withhold for Validation: 0

OK Cancel Delete Transform Help
Adding Forecasts to Data Table

<table>
<thead>
<tr>
<th>Quarter</th>
<th>GDP</th>
<th>Building starts</th>
<th>Housing starts</th>
<th>Permits</th>
</tr>
</thead>
<tbody>
<tr>
<td>from BEA</td>
<td>billions of chained 2009 dollars, seasonally adjusted</td>
<td>from OECD, seasonally adjusted</td>
<td>thousands of units, from FRED</td>
<td>thousands of permits, from FRED</td>
</tr>
<tr>
<td>214</td>
<td>Q2/13</td>
<td>15606.6</td>
<td>72055.6</td>
<td>866</td>
</tr>
<tr>
<td>215</td>
<td>Q3/13</td>
<td>15779.9</td>
<td>73500.0</td>
<td>883</td>
</tr>
<tr>
<td>216</td>
<td>Q4/13</td>
<td>15916.2</td>
<td>85416.7</td>
<td>1012</td>
</tr>
<tr>
<td>217</td>
<td>Q1/14</td>
<td>15831.7</td>
<td>77083.3</td>
<td>934</td>
</tr>
<tr>
<td>218</td>
<td>Q2/14</td>
<td>16010.4</td>
<td>82111.1</td>
<td>984</td>
</tr>
<tr>
<td>219</td>
<td>Q3/14</td>
<td>16205.6</td>
<td>85805.6</td>
<td>1029</td>
</tr>
<tr>
<td>220</td>
<td>Q4/14</td>
<td>16294.7</td>
<td>88555.6</td>
<td>1055</td>
</tr>
<tr>
<td>221</td>
<td>Q1/15</td>
<td>16264.1</td>
<td>80722.2</td>
<td>975</td>
</tr>
<tr>
<td>222</td>
<td>Q2/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>223</td>
<td>Q3/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>224</td>
<td>Q4/15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Add Regressor Variables
Revised Model

Automatic Forecasting - GDP
Data variable: GDP (billions of chained 2009 dollars, seasonally adjusted)

Number of observations = 221
Time indices: Quarter (from BEA)

Forecast Summary
Math adjustment: Square root
Forecast model selected: ARIMA(1,1,0) with constant + 3 regressors
Number of forecasts generated: 3
Number of periods withheld for validation: 0

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Estimation</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>59.9928</td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>42.9407</td>
<td></td>
</tr>
<tr>
<td>MAPE</td>
<td>0.520399</td>
<td></td>
</tr>
<tr>
<td>ME</td>
<td>0.147357</td>
<td></td>
</tr>
<tr>
<td>MPE</td>
<td>-0.00510692</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model Summary

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)</td>
<td>0.126197</td>
<td>0.0688097</td>
<td>1.834</td>
<td>0.068056</td>
</tr>
<tr>
<td>LAG(Permits,1)</td>
<td>0.000963898</td>
<td>0.000202409</td>
<td>4.76213</td>
<td>0.000004</td>
</tr>
<tr>
<td>LAG(Permits,2)</td>
<td>0.000713768</td>
<td>0.000200952</td>
<td>3.55193</td>
<td>0.000471</td>
</tr>
<tr>
<td>LAG(Permits,3)</td>
<td>0.000495638</td>
<td>0.00020176</td>
<td>2.45658</td>
<td>0.014830</td>
</tr>
<tr>
<td>Mean</td>
<td>0.330946</td>
<td>0.0238401</td>
<td>13.8819</td>
<td>0.000000</td>
</tr>
<tr>
<td>Constant</td>
<td>0.289182</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Backforecasting: yes
Estimated white noise variance = 0.0936222 with 212 degrees of freedom
Estimated white noise standard deviation = 0.305977
Number of iterations: 7
Revised Forecasts

Forecast Plot for GDP
ARIMA(2,1,0) with constant
Q1/14 Q1/15 Q1/16
Quarter
16000
16200
16400
16600
16800
17000
GDP
actual
forecast
95.0% limits
Note

• At 8:30AM this morning (June 24) the BEA announced a “third” estimate of the Q1/2015 GDP. It raised the estimate from 16,264.1 to 16,287.7.

• That’s a revision from -0.7% to -0.2% in the annual rate compared to the previous quarter.

• That changes our Q2/2015 forecast from 16,334.8 to 16,361.5. That’s an increase from about 2.0% to 2.2% growth year-over-year.
Data Sources

- New Private Housing Units Authorized by Building Permits – Federal Reserve Bank of St. Louis

- GDP – Bureau of Economic Analysis, U.S. Department of Commerce
References


Recorded Webinar

• You may find the recorded webinar, PowerPoint slides and sample data at:

  www.statgraphics.com

• Look for “Instructional Videos”.