Time Series Basics

Patrick Herb

Brandeis University

Spring 2016
Course Overview

1. Time Series Basics
2. Trends
3. Seasonality
4. Cycles
Get to Know Your Data

- Understand Source
- Understand Units
- Understand Frequency
- Have an idea about the variance structure
Necessary Details

- Frequency of measurement
- Type
  - Time Average (Some Prices)
  - Flow Measure (Consumption)
  - Point Measure (Stock Price)
- Seasonally Adjusted?
- Revised? (When did you know the data?)
Necessary Details

Who Collected the Data?

- Objective Source?
- Biased Source?
  - Source has conflict of interest
  - Analyst Forecast
  - Is there an agenda behind the data?
A Model

- What sort of mechanism could have generated the data?
- Do you want the true data generating process or an approximation?
- Often lacking exact theories
- Some guidance is useful
- Total economic ignorance can be a problem: Black-Box forecasting methods
Step 1: Plot Data

- Always plot your data first!
- Look for:
  - Trends
  - Seasonalities
  - Regime Changes
  - Outliers
Anscombe Data Sets

- Four different \((x,y)\) data sets
- Same regression equation:
  \[ y_i = \beta x_i + \epsilon_i \]  

1. All have same regression output:
   - R-Squared
   - \(SSR = \sum e_i^2\)
   - \(\hat{\beta}\)

2. Data are completely different
Anscombe Data Set

![Graphs of the Anscombe Data Set](image-url)
Trends

- Is data increasing or decreasing over time?
- Linear or Log?
- Trend Breaks
The image shows a graph titled "U.S. Real GDP". The x-axis represents the dates from 1947q3 to 2015q3, while the y-axis represents billions of chained 2009 dollars. The graph illustrates the trend of U.S. real GDP over this period, with a general upward trend indicating growth.
U.S. Real GDP

- Log Real Gross Domestic Product Fitted values

Date

Billions of Chained 2009 Dollars

- 7.5, 8, 8.5, 9, 9.5, 10

- Log Real Gross Domestic Product
- Fitted values

Patrick Herb (Brandeis University)
Fitting a Linear Trend

Code for Stata

```
freduse GDPC1
gen lgdp = log(GDPC1)
tsset daten
gen dateq = qofd(daten)
tset dateq
gen t = _n
```

Compare two regression models

```
reg lgdp t
reg lgdp dateq
```

Why is the constant different?
Creating variables for estimation on subsets of data:

```plaintext
gen month = month(daten)
gen year = year(daten)
gen after89 = year > 1989
gen start90 = year(daten) > 1989
display tm(2008m1)
display tq(2008q1)
gen start89q2 = dateq > tq(1989q1)
```
Fitting a Linear Trend

Try regressing on subset of data

```plaintext
reg lgdp t if year > 1989
reg lgdp t if start90 == 1
reg lgdp t if dateq >= tq(1990q1)
```
Seasonality

Many series have strong seasonal patterns
- Housing Starts
- Gasoline Sales
- Liquor Sales
- Retail Sales

Patterns are usually there for obvious reasons
Housing Starts - Seasonally Adjusted

Thousands of Private Housing Units Started

1960m1 1980m1 2000m1 2020m1
datem

Patrick Herb (Brandeis University)
Housing Starts - Not Seasonally Adjusted

Average Housing Starts by Month

<table>
<thead>
<tr>
<th>Month</th>
<th>Housing Starts (Thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>87.1439</td>
</tr>
<tr>
<td>2</td>
<td>90.1754</td>
</tr>
<tr>
<td>3</td>
<td>120.389</td>
</tr>
<tr>
<td>4</td>
<td>136.781</td>
</tr>
<tr>
<td>5</td>
<td>142.189</td>
</tr>
<tr>
<td>6</td>
<td>141.007</td>
</tr>
<tr>
<td>7</td>
<td>134.977</td>
</tr>
<tr>
<td>8</td>
<td>133.109</td>
</tr>
<tr>
<td>9</td>
<td>126.298</td>
</tr>
<tr>
<td>10</td>
<td>130.179</td>
</tr>
<tr>
<td>11</td>
<td>108.858</td>
</tr>
<tr>
<td>12</td>
<td>92.2596</td>
</tr>
</tbody>
</table>
Housing Starts - Not Seasonally Adjusted

Stata Code for Previous Slide

```stata
freduse HOUSTNSA
tset daten
gen month = month(daten)
graph bar (mean) HOUSTNSA, over(month) blabel(bar) \ 
ytitle(Housing Starts (Thousands)) \ 
title(Average Housing Starts by Month)
```

Patrick Herb (Brandeis University)
Crude Seasonal Adjustment

- Generate Dummy Variables for Months of Year
- Regress on Dummies (No Constant, Why?)

\[ y_t = \sum_{i=1}^{12} \beta_i S_{it} + \epsilon_t \] (2)

- Store Residuals

```
tabulate month, gen(S)
reg HOUSTNSA S*, noconstant
predict e, residual
* annualize
gen e_annual = 12*e
```
How does the government seasonally adjust the data?

- X-13 Filter
- Complicated Filter
- Available for Matlab
Housing Starts: Crude Adjustment

Month

Thousands of Units Started

1960m1 1980m1 2000m1 2020m1

Monthly

Monthly Seasonality Removed Not Seasonally Adjusted

Patrick Herb (Brandeis University)

Time Series Basics

ECON/FIN 250: Spring 2016
Housing Starts: Crude Adjustment

![Graph showing Housing Starts over time with Crude and Government Seasonal Adjustments.]

- **Crude Seasonal Adjustment**
- **Government Seasonal Adjustment**

**Patrick Herb (Brandeis University)**

**Time Series Basics**

**ECON/FIN 250: Spring 2016**

Page 25 of 29
Cycles

- Third Component of Time Series
- Most Difficult to Define
- Typical: Cyclicality
- More General: Any sort of dynamics not captured by trends and cycles
  - Dynamics
  - Persistence
  - Link Between Present, Past and Future
U.S. Unemployment Rate - SA

![Graph of U.S. Unemployment Rate](image)

Civilian Unemployment Rate

Monthly

Patrick Herb (Brandeis University)

Time Series Basics

ECON/FIN 250: Spring 2016
Cycles: Unemployment Rate

- Can you model the exact periodicity of?
- How many cycles are there?
- How can we model this?
- Consider the asymmetries and nonlinearities
Trend, Seasonality, Cycles

Most economics, finance and business data have at least one of these components.

Modeling real world data can be difficult.

Engineering & Science approaches can be less relevant.

Modeler need to think carefully about data and model.