SHOWDOWN AT THE SAPW!

Forecasts versus Published Data

Elijah L. Hood and Catherine C.H. Hood

Catherine Hood Consulting
SAPW, November 2019
Background

- A subset of clients purchase our one-step-ahead forecasts so they can better predict the movements of the stock markets around press releases.

- To help them with their predictions, we do a periodic review of
  - how our forecasts stack up against the first published numbers, and
  - possible residual seasonality in the first published data.
## Data - Anonymous

<table>
<thead>
<tr>
<th>Data Source</th>
<th>No. of Series</th>
<th>Min Length</th>
<th>Max Length</th>
<th>Mode Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long Series</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source A</td>
<td>29</td>
<td>3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Source B</td>
<td>17</td>
<td>8</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Source C</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Short Series</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source D</td>
<td>11</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Source E</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
# Data - US Starts and Permits, and NAR

<table>
<thead>
<tr>
<th>Long Series</th>
<th>Data Source</th>
<th>No. of Series</th>
<th>Min Length</th>
<th>Max Length</th>
<th>Mode Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Housing Starts</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>US Building Permits</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>NAR</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Short Series</td>
<td>none</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Short series analysis

- Finding seasonality through visual inspection
  - Percent differences forecast versus published
  - Percent differences forecast versus published, by month
  - First difference of the published series, by month
Formulas

- Percent differences:

\[ D = 100\% \times (1 - \frac{p}{f}) \]

where

\( p = \) first published seasonally adjusted series

\( f = \) projected/forecasted seasonally adjusted series

- Published first differences: \( p_t - p_{t-1} \)
Visual Inspection: Percent Differences

Source D

Year

Percent Differences

2017 2018 2019 2020
Percent Differences by Month

Source D by Month

Percent Differences

Month
First Difference of Published, by Month
Finding Bias: Formulas

- Let \( D \) = percent differences
- Sum: \( \sum D \)
- Average Percent Differences: \( \sum \frac{D}{n} \)
- Absolute Average Percent Differences: \( \sum \frac{|D|}{n} \)
### Bias Stats Summary Table

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Series</th>
<th>Sum</th>
<th>Average</th>
<th>Absolute Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source C</td>
<td>Shows most bias</td>
<td>10.03</td>
<td>0.39</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>Shows least bias</td>
<td>2.17</td>
<td>0.08</td>
<td>0.73</td>
</tr>
<tr>
<td>Source D</td>
<td>Shows most bias</td>
<td>38.09</td>
<td>1.41</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td>Shows least bias</td>
<td>2.76</td>
<td>0.10</td>
<td>1.84</td>
</tr>
<tr>
<td>Source E</td>
<td>Shows most bias</td>
<td>-12.79</td>
<td>-0.19</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>Shows least bias</td>
<td>-0.19</td>
<td>0.00</td>
<td>0.24</td>
</tr>
</tbody>
</table>
Long Series Analysis

- Seasonality diagnostics from X-13:
  - M7 stats showed no indication of seasonality.
  - Out of 76 long series,
    - 11 had seasonal peaks;
    - 7 had seasonal peaks at frequencies 1, 2, or 4;
    - 4 had seasonal peaks at frequency 5; and
    - 8 had peaks that were almost seasonal but not quite significant.
Seasonal Peaks in Spectrum Graphs

Spectrum of the Prior Adjusted Original Series - Proj vs Pub
Benchmarking Spikes

Original Series - Proj vs Pub

Value

Date

-2.50 -2.25 -2.00 -1.75 -1.50 -1.25 -1.00 -0.75 -0.50 -0.25 0.00 0.25 0.50 0.75

2014 Bias Spikes
An Oddity