Abstract
The business cycle has been a subject of great economic interest over the past
century. Decision making in both the public and private sector is influenced
by the phase of the business cycle, and as a result, our ability to understand
model real economic activity is incredibly important. This study presents
a group of linear models that attempt to explain the evolution of real eco-
nomics activity, in an effort to determine how the inclusion of leading indica-
tors affects out-of-sample predictive power. I focus on 10 potential leading indica-
tors: interest rate spread, producer price index, hours worked, corporate
profits, M1, M2, the Federal Funds Rate, the S&P 500, and the Dow Jones
industrial average. Using a rolling vector autoregressive structure and two
different forecasting methods, all possible combinations of these leading indica-
tors were analyzed. I found that including any of the viable leading indic-
dators in the model improves performance, however interest rate
spread affects out-of-sample predictive power. I focus on 10 potential leading
indicators: interest rate spread, producer price index, hours worked, corporate
profits, M1, M2, the Federal Funds Rate, the S&P 500, and the Dow Jones
industrial average. Using a rolling vector autoregressive structure and two
different forecasting methods, all possible combinations of these leading indica-
tors were analyzed. I found that including any of the viable leading indi-
cator candidates in the model improves performance, however interest rate
spread, the producer price index, and M1 yield the best results. With every
additional variable beyond two included in the regression, the loss in degrees
of freedom results in worse forecasts despite better in-sample fit.

Potential Leading Indicators

<table>
<thead>
<tr>
<th>Corporate Profits</th>
<th>Federal Funds Rate</th>
<th>S&amp;P 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Loans</td>
<td>Interest Rate Term Spread</td>
<td>Dow Jones</td>
</tr>
<tr>
<td>M1 and M2</td>
<td>Producer Price Index</td>
<td>Hours Worked</td>
</tr>
</tbody>
</table>

Preliminary Data Analysis: Unit Root Testing
Unit root testing was performed using the Augmented Dickey Fuller test,
based on a regression of the form:
\[ \Delta y_t = \alpha + \beta t + \delta y_{t-1} + \sum_{i=2}^{l} \gamma_i \Delta y_{t-i} + \epsilon_t \]
Non-stationary sequences were differenced until they became stationary.

Preliminary Data Analysis: Static Granger Causality
Granger causality tests were performed to determine the temporal relation-
ship between variables. The test statistic is calculated by the estimation of:
\[ y_t = a_0 + \sum_{i=1}^n a_i y_{t-i} + \sum_{i=1}^n b_i z_{t-i} + \epsilon_t \]
This test eliminated Hours Worked as a potential leading indicator.

VAR Specification and Lag Selection
Forecasting was performed based on a series of multivariate vector autore-
gressions. These took the standard form:
\[ x_t = \alpha + \sum_{i=1}^{\max} \gamma_i x_{t-i} + u_t \]
For consistency in the number of required parameter estimates, I chose
to use a fixed number of lags across all regressions. To the
left is a plot of selected lag orders for a random draw of regressions
used in the forecasting process, calculated using the Hannon-Quinn criterion. On average
I found that one lag was optimal, regardless of the number of time periods included in the
rolling regression. Therefor, all regressions were calculated using an lmax=1.

Conclusions
- Leading indicators provide valuable forecasting information
- Interest Rate Spread, M1, and PPI produce the best bivariate VAR forecasts.
- Including more than 1 leading indicator worsens model performance