Forecasting Recessions and Bayesian Model Averaging

Jeremy Piger

University of Oregon

University of Oregon Economics Club

November 17, 2014
My Labels

- Empirical Macroeconomist
  - What are the long-run effects of recessions?
  - Can we improve forecasts and nowcasts of recessions?
  - Subnational business cycles
  - The “Great Moderation”

- Applied Econometrician
  - Time Series Econometrics
  - Bayesian Econometrics
Today I will speak about forecasting and nowcasting recessions using Bayesian model averaging.

Most closely related to my paper “Forecasting National Recessions Using State-Level Data.”

Joint work with Michael Owyang (FRB St. Louis) and Howard Wall (Lindenwood University) and forthcoming in the Journal of Money, Credit and Banking.
Motivation:

- We now date the peak of the 2008-2009 “Great Recession” to December 2007.
- However, statistical models being used in real time didn’t send a definitive signal of the start of this recession until late 2008.
- Can’t we do any better?
What econometric models do we use to forecast recessions?

- Define $S_t \in \{0, 1\}$ as a dummy variable indicating whether month $t$ is an expansion or recession month.

- Our objective is to forecast $S_{t+h}$ using predictors available to a forecaster at the end of month $t$. Call these $X_{1,t}$, $X_{2,t}$, $\ldots$, $X_{K,t}$.

- We would like a model that uses the information in the $X$ variables to give us the probability that the economy will be in recession at time $t$: $\Pr(S_{t+h} = 1)$
• Just run a linear regression?

\[ S_{t+h} = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots + \beta_K X_{k,t} + u_t \]

• Might be a bad idea. The fitted value from this regression could be negative or greater than one.

• Instead we use “discrete data” models:

\[ \Pr (S_{t+h} = 1) = \Pr (\beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots + \beta_K X_{k,t} + u_{t+h} > 0) \]

\[ u_{t+h} \sim N(0,1) \]
Forecasting and Nowcasting Recessions

- Forecasting and nowcasting recessions has received significant attention from academics, policymakers, and practitioners.

- A summary of findings:
  - Variables capturing interest rate spreads are the only predictors of recessions at longer horizons (beyond 3-6 months ahead).
  - Variables capturing real economic activity, asset prices, and the level of interest rates have predictive power at shorter horizons.
We investigate whether state-level economic activity can be used to improve monthly forecasts of the U.S. business cycle phase.

Baseline model contains national variables that have been found to forecast business cycle phases.

- Interest rate spreads and levels; equity returns; employment growth; industrial production growth.

State-level economic activity is measured using state-level employment growth.
• There are 50 state-level variables. Which ones should go in the model to produce recession forecasts?

• We could just put them all in. This leads to a model with many parameters and a potentially high level of estimation uncertainty. Won’t forecast well.

• We could try a select a single “best” model. But why focus on only one model?

• Here we use Bayesian Model Averaging (BMA) to average forecasts.
Bayesian Model Averaging

• The Bayesian approach to econometrics treats unknown objects of interest as random variables.

• We then express what we know about these unknown objects of interest using a probability density function.

• This includes the identity of the “true model.”

• By taking a Bayesian approach, we can compute:

\[ \text{Pr (Model is true model|Data)} \]

• This is called a “Posterior Model Probability.”
Bayesian Model Averaging

How do we form the Posterior Model Probability?

Bayes Rule!

\[ Pr(A|B) \propto Pr(B|A) Pr(A) \]

\[ Pr(\text{Model is true model}|\text{Data}) \]

\[ \propto Pr(\text{Data}|\text{Model is true model}) Pr(\text{Model is true model}) . \]
Bayesian Model Averaging

• Suppose we have a forecast of $S_{t+h}$ that comes from a particular model, called $Model_j$. Call this:

$$Pr (S_{t+h} = 1 | Model_j)$$

• Also suppose that there are $N$ possible models - so $j = 1, 2, \ldots, N$.

• Our Bayesian Model Averaged Forecast is:

$$Pr (S_{t+h} = 1) = \sum_{j=1}^{N} Pr (S_{t+h} = 1 | Model_j) Pr (Model_j \text{ is true model} | Data)$$
Out-of-Sample Forecasting Exercise

• Initial estimation period is August 1960-December 1978.

• Out-of-sample forecasts are computed with recursive estimation through June 2011.

• Out-of-sample period contains five NBER recession episodes, accounting for 15% of months.

• Forecasts are produced at the $h = 0, 1, 2,$ and 3 month horizons.
Timeline for Forecasts

September

h=0

October

h=1

November

h=2

December

h=3

September 30
## Out-of-Sample Forecast Evaluation Metrics

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Baseline CSP</th>
<th>Baseline QPS</th>
<th>Extended CSP</th>
<th>Extended QPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 0$</td>
<td>0.92</td>
<td>0.11</td>
<td>0.94</td>
<td>0.08</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>0.91</td>
<td>0.13</td>
<td>0.95</td>
<td>0.09</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.90</td>
<td>0.14</td>
<td>0.90</td>
<td>0.15</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.90</td>
<td>0.15</td>
<td>0.89</td>
<td>0.18</td>
</tr>
</tbody>
</table>
## Out-of-Sample Forecast Evaluation Metrics - Expansion Months

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Baseline</th>
<th>Extended</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CSP</td>
<td>QPS</td>
</tr>
<tr>
<td>$h = 0$</td>
<td>0.96</td>
<td>0.06</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>0.95</td>
<td>0.07</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.95</td>
<td>0.07</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.97</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Out-of-Sample Forecast Evaluation Metrics - Recession Months

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Baseline CSP</th>
<th>Baseline QPS</th>
<th>Extended CSP</th>
<th>Extended QPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 0$</td>
<td>0.68</td>
<td>0.41</td>
<td>0.91</td>
<td>0.15</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>0.64</td>
<td>0.47</td>
<td>0.88</td>
<td>0.23</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.61</td>
<td>0.55</td>
<td>0.61</td>
<td>0.51</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.54</td>
<td>0.64</td>
<td>0.52</td>
<td>0.71</td>
</tr>
</tbody>
</table>
### Forecasting Results - 2008-2009 Recession

<table>
<thead>
<tr>
<th>Date</th>
<th>$S_t$</th>
<th>Baseline (h=1)</th>
<th>Extended (h=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 2007</td>
<td>0</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>December 2007</td>
<td>0</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>January 2008</td>
<td>1</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>February 2008</td>
<td>1</td>
<td>0.36</td>
<td>0.38</td>
</tr>
<tr>
<td>March 2008</td>
<td>1</td>
<td>0.19</td>
<td>0.61</td>
</tr>
<tr>
<td>April 2008</td>
<td>1</td>
<td>0.42</td>
<td>0.90</td>
</tr>
<tr>
<td>May 2008</td>
<td>1</td>
<td>0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>June 2008</td>
<td>1</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>July 2008</td>
<td>1</td>
<td>0.30</td>
<td>0.92</td>
</tr>
<tr>
<td>August 2008</td>
<td>1</td>
<td>0.80</td>
<td>1.00</td>
</tr>
<tr>
<td>September 2008</td>
<td>1</td>
<td>0.76</td>
<td>1.00</td>
</tr>
<tr>
<td>October 2008</td>
<td>1</td>
<td>0.88</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Average Predictor Inclusion Probabilities for Recursive Estimations

\[ h = 1 \]
Conclusion

- We have investigated whether state-level employment data improves short-horizon forecasts of expansion and recession phases.
- We use Bayesian methods to incorporate uncertainty about which states should be included.
- There are substantial forecasting improvements from including state-level data during recession months.
- There is significant uncertainty regarding which states improve forecasts.