“Evaluating Aggregate Survey-Based Indicators in times of crisis.”

ANNA LEACH & ANDREAS BELEGRATIS

14/11/16
I. Brief outline of ITS, the dataset and quantification methods used to provide “early” economic indicators

II. Evaluate their performance in normal times vs crisis using an application on UK manufacturing output.

III. Examine the impact of quantification methods on the performance of survey data in normal times vs financial crisis

IV. Examine the effect of the changes in sample size on the performance of survey data in normal times vs financial crisis

V. Summary
The goal is to:

“Identify and evaluate any significant changes in performance of aggregate survey-based indicators in times of financial crisis.”

We examine two potential effects:
1. Due to quantification methods
2. Due to the changes in sample size

To investigate these effects we proceed with an application on UK’s manufacturing output.
PART I: Application on the ITS

Let's begin with the QITS that asks manufacturing firms the following questions about their volume of output:

- 1] **WHAT WAS THE TREND OVER THE PAST 3-MONTHS??**

<table>
<thead>
<tr>
<th>Possible Answers</th>
<th>Aggregate Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>$U_t$</td>
</tr>
<tr>
<td>Same</td>
<td>$S_t$</td>
</tr>
<tr>
<td>Down</td>
<td>$D_t$</td>
</tr>
</tbody>
</table>

- 2] **WHAT IS THE EXPECTED TREND OVER THE NEXT 3-MONTHS??**

<table>
<thead>
<tr>
<th>Possible Answers</th>
<th>Aggregate Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>$t-1U^e_t,$</td>
</tr>
<tr>
<td>Same</td>
<td>$t-1S^e_t$</td>
</tr>
<tr>
<td>Down</td>
<td>$t-1D^e_t$</td>
</tr>
</tbody>
</table>

- Q1 is considered a retrospective question denoted as “nowcast”
- Q2 is considered a prospective question related to expectations denoted as “forecast”
ITS Publication and Matching with ONS

Q1: is retrospective and corresponds to 2008Q1

Q2: is prospective and corresponds to 2008Q2

ONS publishes official data for 2008Q1 on May (1-month after ITS)

CBI publishes ITS survey data for 2008Q1 on April
Question: How to get from qualitative answers for firms’ expectations to quantitative measures for the output denoted as ($y$)?

Answer: Quantification Methods

1. Balance Method: Anderson (1951) : $y_e^{BAL} := B_t^e = U_t^e - D_t^e$

2. Probability Method: Theil (1952) and Carlson & Parkin (1975)

Based on the idea:

- Each firm has a subjective probability distribution when forming expectations about the volume of output, conditional to the information available at that time.

- Each firm reports “Up” if the mean of that distribution $y_t^e$ is above a threshold interval and reports “Down” if it is below.

- If the mean falls inside the threshold interval $[-b_{it}, a_{it}]$ then it is considered as no growth and the firm reports it will stay the “Same”.

**Figure 1**

- “Up”, if the mean is above $[-b_{it}, a_{it}] \leq y_{i,t}^e$
- “Same”, if the mean is inside $-b_{it} < y_{i,t}^e < a_{it}$
- “Down”, if the mean is below $y_{i,t}^e \leq [-b_{it}, a_{it}]$
Carlson & Parkin (1975) proposed Indicator

- Carlson & Parkin (1975) make 3 critical assumptions:
  1. Firms are independent and follow a common probability distribution $F$.
  2. The common distribution is normal $N(y_t^e, \sigma_t)$.
  3. The threshold interval $[-b_{it}, a_{it}]$ is constant and symmetric $[-\lambda, \lambda]$.

Then by taking the probability of firms expecting “Up” and “Down” and solving for the mean $y_t^e$ one has an explicit formula for an “early” indicator

$$y_{t}^{e CP} := y_{t}^{e} = \lambda \left( \frac{F^{-1}(1 - U_t^e) + F^{-1}(D_t^e)}{F^{-1}(1 - U_t^e) - F^{-1}(D_t^e)} \right)$$
Regression Method

- Based on the idea
  - To use the relationship between firms' perceptions of the past with official realizations as a yardstick to quantify expectations.

Regression Method: Anderson (1952) approach

**Step 1:** Regression between output and retrospective data

1. \( y_t = aU_t - bD_t + u_t \)

**Step 2:** use the OLS estimators \( \hat{a} \) and \( \hat{b} \) to obtain a quantified measure for expectations as an early economic indicator.

2. \( y_{t+1}^{AND} := \hat{a} \ast U_t^e - \hat{b} \ast D_t^e \)
Quick remarks

- Survey data can be used to provide two early indicators before the official data.
- One as a **Forecast** and one as a **Nowcast**.
- Forecast Indicator is given by converting the qualitative expectations into quantitative measures called quantified expectations.
- We outlined 3 conversion approaches to get quantified expectations $y_t^e$: **Balance** ($y^{BAL}$), **Probability** ($y^{CP}$) and **Regression** ($y^{AND}$).
- Next stop measuring the performance of the quantified expectations $y_t^e$ and then comparing between normal times vs financial crisis.
PART II: Measuring the performance of Survey-Based Indicators on Expectations

\[ t-1y_t^e \text{ vs } y_t \]

1. How much they differ in terms of magnitude.

Measure of magnitude:

\[
\text{RMSE}(y_t, y_t^e) = \sqrt{\frac{\sum_{i=1}^{T} (y_t - y_t^e)^2}{T}}
\]

\[
\text{ME}(y_t, y_t^e) = \frac{\sum_{i=1}^{T} (y_t - y_t^e)}{T}
\]

2. How much they are correlated to each other.

Measure of linear relationship:

\[
\text{Correlation}(y_t, y_t^e) = \frac{\sum_{i=1}^{T} (y_t - \bar{y}_t)(y_t^e - \bar{y}_t^e)}{\sqrt{\sum_{i=1}^{T} (y_t - \bar{y}_t)^2 \sum_{i=1}^{T} (y_t^e - \bar{y}_t^e)^2}}
\]

The evaluation of survey based indicators is based on their ability to forecast the official output.
The evaluation of survey data is based on their ability to forecast the output $y$.

The evaluation of forecasting is broken down in two compartments:
1. Evaluation on data we have seen - called train set
2. Evaluation on data we assume unobserved - called test set
3. We split the test set in two periods: a normal period and the financial crisis period until 2009Q3.

Forecasting Experiment:
- We estimate the indicator’s model on the train set and evaluate the indicator’s forecasting performance on the test set.
- To identify changes we need to observe a difference in the indicator’s performance between the normal period and the financial crisis.
A survey based Indicator should perform similarly in normal times and times of crisis this could translate:

\[
\frac{RMSE(y^e,y)_{normal}}{RMSE(y^e,y)_{crisis}} \approx 1
\]

no difference in performance between the two periods

We run the previous experiment on the UK manufacturing output and we found

RESULTS:

1. The Survey-Based Indicator on expectations behaves differently in normal times compared to the financial crisis.

2. The RMSE and ME is considerably lower in normal times

Now we need to try identify the source of that error.
The difference in performance we observed between Normal and Crisis period can be attributed to **3 types of error**:

1. **Expectation error:**
e.g. Firms’ inability to forecast the movements of output.

2. **Conversion error:**
This relates to quantification methods not being able to capture the true values of firms expectations

3. **Sampling error:**
e.g. due to significant changes in sample size affecting the quality of survey data.

Since survey data are approximations of a sample of firms’ expectations, they inevitably contain a **conversion error** and a **sampling error**.

**Question:** Is the conversion or sampling error the reason behind the difference in performance in normal vs crisis?
PART III: The Impact of the conversion error between Normal and Crisis

- In a perfect world one should expect the performance between different indicators be similar inside each period:

  - Normal Times: \( \text{RMSE}^\text{CP}_{\text{normal}} \approx \text{RMSE}^\text{BAL}_{\text{normal}} \approx \text{RMSE}^\text{AND}_{\text{normal}} = a \)

  - Crisis Times: \( \text{RMSE}^\text{CP}_{\text{crisis}} \approx \text{RMSE}^\text{BAL}_{\text{crisis}} \approx \text{RMSE}^\text{AND}_{\text{crisis}} = b \)

- To identify the conversion error we keep everything else constant and we try different quantification methods:

  - We run the experiment
    1. Using different quantification methods: Balance/ Probability/Regression along with extensions of these models.
    2. We also added ARMA models as a benchmark of comparison
RESULTS from the experiments:

1. Survey-Based Indicators produced by many different quantification methods overall perform similarly.

2. Firms could not predict the magnitude of the crisis but could predict the continuous downward trend.

3. Survey-Based Indicators significantly outperform ARMA models in Crisis periods.

4. The latter indicates that firms seem to form expectations rationally in crisis. Firms seem to recognise the continuous downward trend and readjust their expectations.

Now let's carry on to measure the effects of changes in sample size.
PART IV: Examine the impact of changes in sample size between normal times and the financial crisis.

- Dataset: 36 observations ranging from Jan 2007 – Sep 2009.

- We create a dummy variable called “period” to split the data in two groups.

- Group 0 is normal times (blue)

- Group 1 is the financial crisis (red)

- First we test if the change in sample size is significant.

- Second we explore the impact of change in sample on the quality of survey data.
We test the following hypothesis:

- $H_0: \mu_0 - \mu_1 = 5\% \text{change} \quad \text{vs} \quad H_1: \mu_0 - \mu_1 > 5\%$

“Did the sample size change more than 5% “between” normal and the financial crisis period?”

We decided to do:

**Paired $t$ – test for unequal variances between periods: “normal” vs “crisis”**

**RESULTS**

We don’t have evidence to reject the null hypothesis $p.\text{value} = 0.98$.

- The change between normal and crisis period was not significant (less than 5%)
Testing the **effect of change** in sample size on the performance of survey data

- Regression with the forecast error \( y_t - y_t^e \) as a dependent variable.
- **Idea:** “*If the change in sample size has an effect on the quality of data, then we should find a systematic pattern between the forecasting error and the sample size during the different periods.*”

1. Start with the model:
   \[
   \text{forecast.error} \sim \text{intercept} + \text{sample.size} + \text{period} + \text{sample.size.period}
   \]

2. Backward elimination procedure:

3. End up with the model: \( \text{forecast.error} \sim \text{intercept} + \text{period} \)

**RESULTS:**

1. Change in sample size between normal vs crisis does not seem to have any impact on the survey data forecasting performance.
2. The “period” still remains. This confirms there is another source of impact that drives, the difference in performance in normal vs crisis.
Summary

Addressed

1. Impact of the financial crisis of 2008 on the performance of survey data
2. The impact many quantification methods on the performance of survey data between normal vs crisis
3. The change in sample size and the impact on the performance of survey data between normal vs crisis.

Results

1. Survey Based Indicators behave differently in Normal vs Crisis.
2. Quantification error is not the primal reason behind the difference.
3. Survey Based Indicators outperform benchmark ARMA models
4. Firms seem to form rational expectations in crisis
5. Change in sample size doesn’t play a significant role in performance of survey data between normal vs the financial crisis period.

Stylized Facts

1. ITS publication advantage over official figures.
2. Two “early” aggregate economic indicators
3. On expectations as a forecast: $y_{t-1}$
4. On past perceptions as a nowcast
5. Three different quantification approaches to provide average quantitative measures as economic indicators
   - $y^{BAL}_t, y^{CP}_t, y^{AND}_t$

Future research

• Testing the ability of quantified survey expectations not as indicators but as predictors.
• Evaluating Disaggregate Survey Based Indicators in times of crisis—panel analysis approach on firm-level data.
Thank you for your attention

Andreas Belegratis