Fed Manufacturing Surveys Provide Insight into National Economy

by Emily Kerr, Pia Orrenius, Jack Wang and Jesús Cañas

Several regional Federal Reserve Banks conduct monthly surveys of manufacturing activity in their districts—New York, Philadelphia, Richmond, Kansas City and Dallas, which produces the Texas Manufacturing Outlook Survey (TMOS). Survey responses are used to construct diffusion indexes, with positive readings typically indicating expansion and negative readings contraction. While each Fed asks slightly different questions, all include a measure of business activity, new orders, shipments and employment.

Although these surveys are intended to measure local conditions and are not designed to track the national economy, their results are correlated with U.S. indicators. Market watchers and policymakers often rely on these monthly regional manufacturing surveys for early insight into important national measures such as industrial production, employment growth and the Institute for Supply Management’s (ISM) manufacturing PMI index.

Using the regional banks’ results to gauge national activity prompts the question: Which survey and which indexes within each survey are most correlated with national indicators and best forecast their behavior?

The recent performances of the surveys and their component indexes were compared using statistical analysis. While several stand out, the Dallas Fed’s TMOS performs well explaining national indicator variations and forecasting the ISM manufacturing index and industrial production.

Fed Districts’ Unique Insights

All five Fed manufacturing surveys provide timely and relevant information about their respective regional economies, but they differ in the extent to which they reflect national economic activity (Chart 1). The composition of manufacturing within each region is unique and not necessarily reflective of the nation as a whole.

The Dallas Fed district, which covers mostly Texas, differs from the nation because of the energy industry’s large presence—not just significant oil and gas production but also refining and petrochemical industries. Thus, the manufacturing sector is disproportionately weighted toward petroleum and petrochemicals manufacturing.

The New York Fed district—largely the state of New York—is not only home to the nation’s biggest banks and financial services companies but also to a sizable share of the U.S. pharmaceutical industry, which is reflected in the composition of its manufacturing sector. The Philadelphia Fed district, covering eastern Pennsylvania, southern New Jersey and Delaware, also has outsized pharmaceutical manufacturing as well as chemical manufacturing.
The Richmond Fed district—Virginia, West Virginia, Maryland, North and South Carolina—has a large presence of food manufacturing. The Kansas City Fed district—mainly Kansas, Oklahoma, Colorado, Nebraska and Wyoming—is home to relatively large shares of the nation’s aerospace manufacturing and oil and natural gas-related machinery manufacturing.

Survey content and methodology also vary among the Feds. The Philadelphia Fed initiated its Business Outlook Survey in 1968 and the others followed. All ask firms about the number of employees, but Richmond and Dallas also inquire about wages, and only Dallas and Kansas City query about the volume of production.

Although the survey reference periods are all the current month, data collection periods vary—New York, Philadelphia and Richmond collect information in the first half of the month, Kansas City and Dallas collect in the latter half. There is also some variation in sample size; the Dallas Fed survey has the largest number of monthly responses, roughly 110, compared with 70 to 100 for the others.

Choosing an index that all the Fed surveys have in common poses a challenge. Cross-survey comparisons typically rely on either a general business activity (GBA) index or a composite index. The ISM manufacturing index is widely used to forecast U.S. gross domestic product (GDP) growth. Given that the Fed manufacturing surveys come out before this index (Chart 2), what insight might they offer regarding upcoming ISM results (which in turn provide a sense of GDP)?

The performance of the Feds’ headline indexes relative to the ISM index over two time periods is shown in Table 1, using correlation statistics. The longer period begins in 2004 and covers a time during which all the Fed surveys were published. The shorter period focuses on the economic recovery period following the Great Recession.

All surveys do better in the longer time period, and their performances are fairly similar. That said, the Philadelphia and Richmond Fed surveys have the highest correlation with the ISM index during both periods. A correlation coefficient of 0.85 means that (0.85)^2, or 72 percent, of the variation in the ISM manufacturing index can be explained by, for example, the headline Philadelphia Fed GBA index. If the correlation is 0.67, then 45 percent of the variation can be explained.

Besides the ISM manufacturing index, there are other barometers of national economic conditions against which to measure the regional surveys. Correlations of the Fed surveys with growth in industrial production and U.S. payroll employment are shown in Table 2. The Richmond Fed survey has the strongest correlation with industrial production growth; the others are clustered about 12 points below Richmond. Meanwhile, the Dallas Fed’s higher correlation with U.S. employment growth edges ahead of the other surveys.

The general business activity indexes are published by the Dallas, New York and Philadelphia Feds. The Richmond and Kansas City Feds, meanwhile, publish composite indexes. Using such top-line results may be practical, but they may not be the best measures of national economic indicators.

**Correlation with ISM Index**

The ISM manufacturing index is widely used to forecast U.S. gross domestic product (GDP) growth. Given that the Fed manufacturing surveys come out before this index (Chart 2), what insight might they offer regarding upcoming ISM results (which in turn provide a sense of GDP)?

The performance of the Feds’ headline indexes relative to the ISM index over two time periods is shown in Table 1, using correlation statistics. The longer period begins in 2004 and covers a time during which all the Fed surveys were published. The shorter period focuses on the economic recovery period following the Great Recession.

All surveys do better in the longer time period, and their performances are fairly similar. That said, the Philadelphia and Richmond Fed surveys have the highest correlation with the ISM index during both periods. A correlation coefficient of 0.85 means that (0.85)^2, or 72 percent, of the variation in the ISM manufacturing index can be explained by, for example, the headline Philadelphia Fed GBA index. If the correlation is 0.67, then 45 percent of the variation can be explained.

Besides the ISM manufacturing index, there are other barometers of national economic conditions against which to measure the regional surveys. Correlations of the Fed surveys with growth in industrial production and U.S. payroll employment are shown in Table 2. The Richmond Fed survey has the strongest correlation with industrial production growth; the others are clustered about 12 points below Richmond. Meanwhile, the Dallas Fed’s higher correlation with U.S. employment growth edges ahead of the other surveys.

**Better-Performing Survey Indexes**

The general business activity indexes attract most attention as measures of regional Fed survey performance, even though they are not always the best-
performing index, particularly not for the Dallas Fed.

For example, the Dallas Fed’s growth rate of orders index has a correlation of 0.55 with industrial production growth and 0.77 with employment growth. While these correlations don’t affect the Dallas Fed’s ranking in Table 2, they are much higher than the correlations with general business activity.

Using regression analysis, the statistical fit of all Dallas Fed component indexes, including general business activity, was compared by seeing how well they could explain variation in the ISM manufacturing index, industrial production growth and payroll employment growth. The Dallas Fed’s growth rate of orders was the best-performing index. It consistently explained more variation in the three national measures than the other indexes did. The same exercise was performed for all the Fed manufacturing surveys, identifying those indexes with the highest explanatory power vis-à-vis the three national indicators. These indexes are used in the analysis below.

**Goodness of Fit Comparisons**

Correlation statistics are a simple and intuitive measure of the relationship between the survey indexes and national indicators. However, they don’t reflect the statistical value added of the Fed surveys as predictors of the national indicators. For example, it might be that the ISM index is well explained by the series’ own past performance and that even highly correlated Fed survey indexes do not provide much additional information.

The Fed surveys were tested by running three multiple regressions (ISM manufacturing index, industrial production growth and payroll employment growth) on the best-performing index from each Fed survey. The results show the Richmond and Dallas Fed surveys provided statistically significant explanatory power for all three national indicators (Table 3). In this case, statistical significance of a survey index means the probability that the coefficient is not different from zero is less than 5 percent (denoted by **) or 10 percent (denoted by *).

In other words, the surveys with statistically significant coefficients are useful in explaining movements in the national indicators. In addition to the Richmond and Dallas Feds, the Philadelphia Fed contributes statistically significant information regarding the ISM index, and Kansas City to employment growth. The R-squared statistic in Table 3 shows the proportion of the variation in the national indicator that the five regional survey measures explain in each regression.

**Recent Forecast Performance**

An additional gauge of the Feds’ surveys’ predictive power vis-à-vis national indicators is their forecast performance. The national indicator was regressed on each Fed survey (using the best-performing index) and three past data points (or lags). The forecast evaluation period ran from July 2011 to June 2014. Each month
during this period, individual Fed survey indexes were used to forecast the value for the national indicator for that same month.

The root mean squared forecast error (RMSFE) measures the squared difference between forecast and actual results. To make the forecast comparisons easier, the forecasting performance of the various Fed surveys is benchmarked against the RMSFE of a model with only lags and no regional Fed survey. Relative RMSFEs are presented in Table 4. Values less than 1 mean that the regional Fed survey data help improve the accuracy of the forecasts; the lower the RMSFE, the more accurate the forecast.

Using this method, the Dallas Fed survey is the most accurate in forecasting industrial production growth because it has the lowest RMSFE. The Dallas Fed survey is ranked second for forecasting the ISM manufacturing index. In the case of payroll employment growth, the use of regional Fed survey data does not result in a better forecast over this period—the basic model with three lags of payroll employment growth outperformed the indexes.

Providing Insight

An analysis of the regional Feds’ manufacturing surveys suggests they can provide important insight into national economic conditions. In recent years, the Philadelphia Fed’s Business Outlook Survey has performed best in forecasting the ISM manufacturing index, followed by the Dallas Fed’s TMOS. TMOS does best in forecasting industrial production; the Richmond Fed survey comes in second.

Still, the appropriate performance measure for each regional Fed survey is the extent to which each reflects economic activity in its district; after all, none of the regional Fed manufacturing surveys were designed to track the national economy.

Nonetheless, analysts and policymakers are eager to use these data to glean insights into the national economy. To accomplish that, analysts would do well to select the best-performing indexes from each survey, instead of the top-line, general business activity measures. Moreover, while it’s tempting to employ correlation statistics to indicate how well two measurements are related (such as the ISM manufacturing index and a particular Fed’s GBA index), these comparisons may say little about a survey’s value added.

Kerr is a business economist, Orrenius a vice president, Wang a senior programmer/analyst and Carías a business economist in the Research Department of the Federal Reserve Bank of Dallas. The authors thank Alexander Chudik and Anil Kumar for their assistance.

Notes

1 The Dallas Fed only surveys firms in Texas, the New York Fed only ones in New York state.
2 At the Dallas Fed, production (not GBA) is the headline index in TMOS because it is considered the best measure of Texas manufacturing output.
3 A correlation statistic of 1 implies a perfect correlation, 0 no correlation.
4 The Dallas Fed TMOS was started in June 2004, last among the existing regional Fed manufacturing surveys.
5 Regressions also included three lags of the dependent variable.

### Table 4: Forecasting National Indicators Using Regional Fed Manufacturing Surveys

<table>
<thead>
<tr>
<th>ISM Manufacturing index</th>
<th>Industrial Production Growth</th>
<th>Payroll Employment Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional survey measure</td>
<td>Relative RMSFE</td>
<td>Regional survey measure</td>
</tr>
<tr>
<td>Philadelphia Fed GBAI</td>
<td>0.92</td>
<td>Dallas Fed growth of orders</td>
</tr>
<tr>
<td>Dallas Fed growth of orders</td>
<td>0.94</td>
<td>Richmond Fed employment</td>
</tr>
<tr>
<td>Kansas City Fed new orders</td>
<td>0.97</td>
<td>Philadelphia Fed employment</td>
</tr>
<tr>
<td>New York Fed GBAI</td>
<td>0.99</td>
<td>Kansas City Fed employment</td>
</tr>
<tr>
<td>Richmond Fed employment</td>
<td>1.02</td>
<td>New York Fed employment</td>
</tr>
</tbody>
</table>

NOTES: A lower relative root mean squared forecast error (RMSFE) indicates better forecasting performance. The baseline model is one with three lags of the national indicator and no regional Fed survey measure. The sample period is June 2004 to May 2014; forecasts run from July 2011 to June 2014. Each entry represents a separate regression and all include three lags of the dependent variable (the national economic measure). Real-time data were used for payroll employment growth. GBAI = general business activity index.