Accurately Predicting Earnings Surprise Through Sharp Consensus Estimates

Eric Morlot, Senior Vice President, FactSet Estimates
Richard Dutheil, Vice President, FactSet Estimates Collection Engineer
Matt Sakey, Vice President, FactSet Estimates Contributor Relations
Xavier Taureau, Vice President, FactSet Estimates Product Line Manager
Accurately Predicting Earnings Surprise Through Sharp Consensus Estimates

Contents
Goal .................................................................................................................................................. 2
Methodology ....................................................................................................................................... 2
Limitations and Universe .................................................................................................................. 3
Rules .................................................................................................................................................. 3
Accuracy and Coverage ..................................................................................................................... 7
History Trend ..................................................................................................................................... 8
Summary ........................................................................................................................................... 9

Goal
To increase the accuracy of the FactSet Estimates database we have added data that enhances analytical capabilities for the user, called Sharp Consensus. Our goal is to define a methodology that provides a more accurate result than the default consensus figure (where opportunities exist). Used in conjunction with the consensus estimate, the Sharp Consensus is an excellent indicator for identifying earnings surprises.

Methodology
Our approach is to detect informal events in revision patterns and to calculate a consensus based on estimates made after the given event. An informal event may be an event relative to the company itself (e.g., profit warning, new contract), relative to the sector (e.g., a significant price change for commodities that will affect the Energy sector), or a natural disaster (e.g., storm, earthquake).

The Sharp Consensus is calculated by identifying a custom window within the 100-day default consensus. This window is calculated by an algorithm that analyzes revision patterns among covering analysts. These revision patterns suggest that new information has entered the market and affects the security’s consensus value.

The date on which a significant number of broker revisions occur is referred to as the Sharp Event Date. This date represents the starting point of the Sharp Consensus window. Sharp Event Dates are determined by multiple broker revisions made within a short time frame and in the same direction.

Once an event is identified, the Sharp Consensus is calculated using all post-event estimates up to the perspective date (by default, “now”). If no Sharp Event Date can be determined within a 100-day window, no Sharp Consensus will be calculated.
Diagram and Concept Definition

- The Sharp Revision Area is the range of dates where a valid group of revisions has been identified.
- The Sharp Event Date is the start date of the Sharp Revision Area. From this date, new analyst revisions are published that, in theory, incorporate the new information that has entered the market.
- The Sharp Consensus window begins at the Sharp Event Date and includes revisions up to the perspective date, “now.”
- All estimates provided in the Sharp Consensus window are used to calculate the Sharp Consensus Estimate.

Limitations and Universe

While we are confident that additional data items can be calculated, the Sharp Consensus is currently calculated only for EPS and Sales for FQ1 and FY1. It is possible to backtest the Sharp Consensus Estimate for the entire FactSet Estimates history: 1994 for France, 1997 for the rest of Europe, and 2000 for the U.S., Canada, and Asia/Pacific.

Rules

FactSet validates the following rules to calculate a Sharp Consensus Estimate.

- **Coverage:** Our goal is to identify a minimum relevant analyst coverage group enabling the algorithm to highlight a revision pattern among the set of analysts covering the stock. In order for the algorithm to accurately identify a revision pattern there must be a minimum number of brokers covering the security. Through backtesting we have identified a broker coverage requirement (coverage threshold) to accurately predict a revision pattern. Should FactSet increase our coverage threshold, there are small opportunities to improve accuracy; however, that change also has a meaningful impact on coverage universe and, as a result, it significantly reduces the value of the Sharp Consensus. Alternatively, we find that reducing the coverage threshold expands our coverage universe, but accuracy and consistency suffer.
To illustrate, the first chart below shows security coverage universe (Y Axis) relative to the number of analysts covering the respective securities (X Axis).

The second chart shares details on the coverage universe and broker coverage relative to the accuracy rate. While analyst coverage and security universe appear to have an inverse relationship (depicted in both charts), the second chart also shows the correlation between analyst coverage and accuracy rate. The figures to the left of the Y Axis show the impact on coverage universe in percent terms, while those on the right side of the chart provide the percent change in accuracy.

- **Contributors**: The Sharp Consensus calculation may include the same set of estimate contributors as the consensus item. This means we consider all estimates within the default 100-day window. Only the latest data point for each broker is taken into account.

- **Time Period**: We look for a set of revisions that occurred over several consecutive business days. We begin by finding the most recent valid estimate revision (default consensus class), and then search back in time for estimate revisions published within the identified time period between two consecutive revisions.
• **Revision Change:** The identified group must contain a minimum number of revisions with a specified change threshold to be valid. A change threshold represents the difference (in percent terms) between the current estimate and the revised estimate. If the updated items are simply confirmations of existing data, it is unlikely that the event will lead to a surprise.

• **Trend (Conviction Ratio):** A large majority of the revisions must follow the same trend to validate that the revisions are the result of new information in the market. To calculate the conviction ratio, we compare upward revisions to downward revisions. The larger group is divided by the total number of revisions. The ratio must meet a minimum threshold, defined by FactSet, in order to qualify as a Sharp Consensus data point.

The chart below depicts a comparison of coverage universe relative to accuracy impact. Accuracy impact represents the trend when the conviction ratio minimum threshold is altered. For example, when the conviction ratio threshold is lowered, we realize a positive impact on the coverage universe (measured in percent terms to the left of the table) and a negative impact on accuracy (measured in percent terms to the right of the table). As the conviction ratio threshold increases, the impact of the coverage universe drops but accuracy increases.

• **Consistency:** The Sharp Consensus must be consistent with the direction of the revisions. Assuming a group of revisions is the result of new information in the market, the direction of the new consensus data item should also be consistent compared to the detail items. If the new consensus item results in a revision moving in the opposite direction of the majority of the contributors’ revisions used to determine the Sharp Revision Area, the number is discarded.

• **Revisions:** When identifying revisions, the algorithm will discard any invalid set and look for another going further back in time, provided the data remains within the 100-day consensus window. If no set matches the rules, the Sharp Consensus Estimate is N/A.

Once a valid Sharp Event Date is identified, the Sharp Consensus is calculated. The window of the Sharp Consensus is from the Sharp Event Date to the perspective date (up to 100 days).

If a group of revisions passes the above criteria, its start date (the day on which the first revision within the Sharp Revision Area appears) and the requested observation date (any day following the Sharp Event Date up to the 100-day consensus) are used as boundaries for the dynamic window. A mean consensus is then calculated for the defined Sharp Consensus period.

**Negative Revision-Specific Rules**

When a majority of brokers revise estimates downward, our research shows that the above rules are not sufficient to accurately predict the surprise.
In addition to the rules described above, the following rules are also applied to determine if the Sharp Event date identified is valid for downward revisions.

Like the upward revision, a negative Sharp Consensus is calculated only if the vast majority of contributors are revising downward, a minimum number of revisions are present, and the revisions cross the required change threshold.

For downward revisions, as the number of analysts covering a given security increases, a larger group of revisions are required to validate the Sharp Consensus. As a result, a negative Sharp Consensus revision on a more broadly covered company requires a greater signal to reflect an informal event.

**Revision Size (EPS and Sales)**

*This rule is exclusively for downward revisions as tests showed no positive impact when applying it for upward revisions.*

For a group of downward revisions to be deemed relevant and trigger a Sharp Consensus calculation, the average size of the revision must surpass an identified proprietary threshold, measured in percent change of the revision versus the consensus. Additionally, the 100-day consensus window is separated into multiple segments. Each segment represents a new threshold where a new revision ‘minimum requirement’ exists. To continue for the full 100-day window, the revision must meet or surpass each respective proprietary threshold. Finally, both EPS and Sales calculations are subject to separate thresholds.

The following chart depicts a comparison of coverage universe (measured in percent terms to the left of the chart) relative to the accuracy gain of negative revisions. Accuracy gain represents the accuracy trend (measured in percent terms to the right of the chart) relative to the size of the group of revising analysts (as the size of the group increases, so does the accuracy rate).

![Chart showing coverage decrease and accuracy gain](chart.png)

The final rule for downward revisions defines the window size of the Sharp Revision Area (length in days). By adding these adjustments, the Sharp Consensus algorithm is able to more accurately track negative signals and predict a negative surprise. Our research confirms the need for defining a maximum length for the Sharp Revision Area for negative revisions. If a revision area exceeds the maximum number of days required within the Sharp Revision Area, the revision cluster does not provide a sufficient signal that an informal event has taken place to impact estimate forecasts, and, as a result, does not meet the Sharp Consensus requirements.
The chart below shows the accuracy trend of negative revisions relative to number of days within the Sharp Revision Area. As frequent revisions continue and extend the number of days within the revision area, the accuracy of negative revisions falls significantly.

**Accuracy and Coverage**

To verify accuracy of the Sharp Consensus model, 20 quarters and 10 years of historical data were backtested on companies included in the Russell 3000 Index. Again, it is possible to backtest Sharp Consensus for the entire FactSet Estimates history.

To define the coverage and accuracy rate, the Sharp Consensus Estimate is calculated one day before the publication of each period and compared to the regular 100-day consensus.

A Sharp Consensus is considered accurate when it predicts the direction of the surprise.

<table>
<thead>
<tr>
<th>Accuracy Average</th>
<th>Russell 3000</th>
<th>MSCI Europe</th>
<th>MSCI Asia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPS - Last 20 Quarters</strong></td>
<td>78%</td>
<td>71%</td>
<td>87%</td>
</tr>
<tr>
<td><strong>EPS - Last 10 Years</strong></td>
<td>72%</td>
<td>72%</td>
<td>74%</td>
</tr>
<tr>
<td><strong>Sales - Last 20 Quarters</strong></td>
<td>79%</td>
<td>75%</td>
<td>88%</td>
</tr>
<tr>
<td><strong>Sales - Last 10 Years</strong></td>
<td>79%</td>
<td>77%</td>
<td>80%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average Number of Companies Covered</th>
<th>Russell 3000</th>
<th>MSCI Europe</th>
<th>MSCI Asia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPS - Last 20 Quarters</strong></td>
<td>290</td>
<td>46</td>
<td>28</td>
</tr>
<tr>
<td><strong>EPS - Last 10 Years</strong></td>
<td>216</td>
<td>243</td>
<td>190</td>
</tr>
<tr>
<td><strong>Sales - Last 20 Quarters</strong></td>
<td>201</td>
<td>51</td>
<td>58</td>
</tr>
<tr>
<td><strong>Sales - Last 10 Years</strong></td>
<td>82</td>
<td>136</td>
<td>109</td>
</tr>
</tbody>
</table>
**History Trend**

The backtest shows that the Sharp Consensus consistently provides more accurate results versus the mean consensus (always above 50%) for both positive and negative algorithms.

The following chart shows the accuracy (%) trend for the Russell 3000 companies having a Sharp Consensus in the last 20 quarters. The black curve shows the general accuracy, the green curve shows the positive model accuracy, and the yellow curve shows the negative model accuracy.

**Accuracy Over Time**

The next chart depicts the change in coverage universe of the Russell 3000 index over time (20 quarters). The two algorithms have an inverse relationship during recessionary and recovery periods. During recessionary periods, a decrease in the coverage universe for positive revisions and an increase in coverage universe for negative revisions is expected behavior. This observation validates that the coverage universes for the two algorithms react to market trends.
Summary

While many methods can be used to improve consensus estimate accuracy, tests using FactSet Estimates data show a high correlation between specific revision patterns and the direction of surprise. For accurate results, the Sharp Consensus seeks to capture the most current data within the default consensus window. The model leverages an algorithm that identifies a start date of the Sharp Consensus window (Sharp Event Date) as the point at which a material event has occurred or new information was made public. As a result, the Sharp Consensus includes detail estimates calculated using the most current market information.

By comparing the Sharp Consensus to the default consensus, we can more accurately predict the direction of a surprise. We can also expect a more accurate estimate when comparing the Sharp Consensus relative to the default consensus.

By incorporating the Sharp Consensus within default FactSet reports including All Estimates, Broker Outlook and Estimate History, end users are able to more accurately predict the direction of the surprise, draw conclusions as to which analysts are incorporating current information within their models, and are attune to events affecting EPS and sales forecasts for a given company.