Why Should I Trust Your Forecasts?

Featuring:
Paul Goodwin, Professor of Management Science
University of Bath, UK
Table of Contents

Introduction ................................................................. 1
What Is Reasonable to Expect of a Forecast? ......................... 1
How Can You Generate Trust in Your Forecasts? ................... 3
   Building Trust – Rapport, Goodwill and Shared Values .......... 3
   Building Trust – Transparency in the Methodology ............. 3
   Building Trust – Candor About Accuracy .......................... 4
   Building Trust – Candor About Uncertainty ...................... 4
   Building Trust – Competence ....................................... 6
Balancing the Drivers of Trust ........................................... 7
Best Practice Recommendations ........................................ 7
   Benchmark Against Other Methods to Show Relative Accuracy 7
   Test the Robustness of the Forecasting Methodology .......... 8
   Average the Results of Multiple Forecasting Methods .......... 8
   Draw on the Talents of Both Quantitative and Business Specialists 8
Closing Thoughts ............................................................ 9
About the Presenter .......................................................... 9
For More Information ...................................................... 9
Introduction

British Gas supplies energy and home services to some 12 million customers in the UK. The utility needs accurate forecasts of daily demand for natural gas so that it can efficiently manage storage, handling and distribution processes.

Some years ago, the company’s mathematicians developed a sophisticated but complex methodology that generally delivered highly accurate forecasts. The model incorporates good logic – it accounts for such factors as daily temperatures and learning from previous forecast errors – but it was too complicated for business users to comprehend. As a result, the shift staff who were supposed to use the forecasts were reluctant to accept them.

“Despite the accuracy of the model, the mathematicians had difficulty selling it to them,” said Paul Goodwin, Professor of Management Science at the University of Bath. “They could not understand the model in qualitative or intuitive terms. They saw it as mathematical gobbledygook. Despite measures presented to them showing the general high accuracy of the model, the staff instead recalled the occasional high errors on single days – and they developed the perception that the method was inaccurate. In short, they did not trust the method, so they did not trust its forecasts.

“An untrusted forecast is a worthless forecast, no matter how accurate it is. If decision makers refuse to use our forecasts, no matter how sophisticated and accurate they might be, we are basically wasting our time. If a forecast is untrusted, people will try to adjust it, override it or totally ignore it. And if they adjust it, all sorts of judgmental biases are likely to reduce its accuracy.”

Paul Goodwin
Professor of Management Science,
University of Bath

Fostering trust is therefore a crucial aspect of forecasting. That was the topic of a webinar sponsored by SAS and Foresight that sought to answer six questions related to trust in forecasts. What is trust? What characterizes a trusted forecast? How many so-called forecasts don’t meet those criteria? When can trust be misplaced? What drives trust? How can you foster trust in forecasts?

What Is Reasonable to Expect of a Forecast?

According to the Oxford English Dictionary, trust is “firm belief in the reliability, truth or ability of someone or something.” This definition can be troublesome for forecasting, because we live in an inherently unpredictable world where perfectly reliable forecasts are not achievable, no matter what senior managers might think. A forecast will seldom equal the true outcome, so forecasts may be seen as inherently untrustworthy.

While it is not reasonable to expect perfect predictions, a trusted forecast can be expected to meet these qualifications:

- Make the best use of available information, subject to cost-effectiveness constraints.
- Be based on the most appropriate forecasting method, correctly applied.
• Be based on reasonable and clearly stated assumptions that can be challenged
  and adjusted as needed.
• Be free of political or motivational biases.
• Honestly represent the uncertainty associated with the forecast.

If these criteria are met, the result will be a competent and honest prediction of future
demand, or whatever else is being forecast. It sounds simple enough, but several types
of so-called forecasts fail these tests of trust and masquerade as forecasts. Here are a
few culprits:

**Projections with an agenda.** Goodwin related a case where an Internet service
provider made uncommonly low estimates of the number of new customers expected
to sign up each week. Week after week, the forecasts were too low. “When I queried
this, I was told that the forecasts were produced by the marketing department, which
looked good if the number of new customers exceeded their forecasts,” said Goodwin.
“Clearly these were not trustworthy forecasts at all; they were not even forecasts in my
book. They were a device for maximizing the political standing of the marketeers. If the
forecasts were initially trusted, I suspect that trust would soon be lost.”

Goodwin described another case where forecasters were tasked to predict the number
of complaints that would be received in a year, so the company could plan for the
workload in their complaints procedure. “Whatever the people forecast, their managers
told them to halve the forecast. So if they forecast 1,000 complaints, he told them to
make it 500, because that number was more politically acceptable.”

**Targets.** “Even if they have the name ‘forecast’ slapped on them, targets are not
trustworthy forecasts,” Goodwin said. “Targets are used to motivate people or state
what we hope to achieve, not what we generally think we will achieve. Indeed, targets
may even be deliberately set unduly high – at the top end of the range of possibilities –
to try to motivate people to achieve more. So targets shouldn’t be trusted.”

**Attempts to influence the outcome.** “Don’t trust forecasts that attempt to influence
the future rather than try to predict how the future will actually turn out,” said Goodwin.
Be wary of an electric car manufacturer’s claim that oil prices will double in the next 10
years. Or the environmental lobbyist’s prediction that world temperatures will rise by 9
F by 2090. Or the consumer electronic manufacturer’s claim that in five years, everyone
will have the company’s flagship product – except you, unless you run out and buy it.

**Expert opinions.** We can be too trusting of statements put forth by industry experts,
since the accuracy of their forecasts varies quite a bit, Goodwin noted. Philip Tetlock,
a psychology professor at the University of Pennsylvania, wrote about a 20-year study
in which 284 experts in many fields were asked to make 28,000 predictions about the
future. The predictions of these government officials, professors, journalists and other
subject-matter experts were only slightly more accurate than chance, and worse than
basic computer algorithms.

“If you are doctoring the forecast – for instance, if you’re pretending losses or complaints
are less likely than they actually are – then the forecast is not very trustworthy. But that’s
easy to say; I haven’t got senior managers breathing down my neck.”

Paul Goodwin
Professor of Management Science,
University of Bath
On the other hand, veteran product managers who know their markets well have a much more accurate view of what future sales are likely to be. Other research showed that professionals in certain fields – such as weather forecasting, accounting and insurance – tended to exhibit good judgment in their forecasts, while those in other fields – such as personnel selectors, student admissions officers and stock brokers – tended to exhibit relatively poor judgment.

**Decisions based on forecasts.** A company might forecast and genuinely expect demand of 200 units for a product, and then decide to stock 240 units to be ready for above-average demand. “The first number – the 200 – is a forecast,” said Goodwin. “The second – the decision to hold stock of 240 units – is a decision. But one retailer I did some research for referred to this latter number as a forecast. They would have said, ‘We forecast 240 units for the next month.’ This will only cause confusion. The decision will almost always be too high when compared to actual demand, and trust will be eroded.”

**How Can You Generate Trust in Your Forecasts?**

The British Gas example shows that accuracy and technical sophistication of a forecasting method are not enough to guarantee trust. So what does influence trust in forecasts? The answer is intuitive, and research has confirmed it: People will trust forecasts more if they believe the forecasters are driven by goodwill and shared aims, are seen to be honest and competent, and have a clear rationale behind their methodology.

**Building Trust – Rapport, Goodwill and Shared Values**

Establish good close relationships with users, and have open communications. Reassure users that you share their motives and objectives. For example, show that you understand the implications of the forecasts on their business objectives, and you share their interests in using the forecast for decision making.

“Use language users are comfortable with,” said Goodwin. “Sometimes the users are mathematicians, but if they’re not, avoid jargon and forecaster-speak. No alphas, betas, gammas or omegas. Otherwise they might start to think you’re a boffin, ignorant of the real world and promoting a technical agenda for its own sake.”

**Building Trust – Transparency in the Methodology**

As much as possible, make your methods intuitively understandable so that you can explain them in nontechnical language. “Avoid walls of numbers, which are probably impenetrable to light and insights and will only confuse and turn off the potential user,” said Goodwin. “Narratives might help – stories of what you think is influencing the forecast – but avoid the narrative fallacy where you find yourself inventing stories to explain what were really random patterns in the past.”

People will trust forecasts more if they believe the forecasters are driven by goodwill and shared aims, are seen to be honest and competent, and have a clear rationale behind their methodology.

“Avoid jargon and forecaster-speak. No alphas, betas, gammas or omegas. Otherwise they might start to think you’re a boffin, ignorant of the real world and promoting a technical agenda for its own sake.”

Paul Goodwin
Professor of Management Science,
University of Bath
There is a danger in being transparent about your forecasting methods. It’s like pulling back the curtain on the great and powerful Wizard of Oz. Sometimes an embarrassingly simple forecasting method performs better than something more sophisticated and situation-aware. “Simple time-series methods that blindly extrapolate past patterns into the future often give more accurate forecasts than explanatory methods, such as multiple regression,” said Goodwin. “But managers might expect your models to provide causal explanations. They might say, ‘Surely it was the price increase that hit our sales last month,’ or ‘Our sales benefited from our new advertising strategy; your models ignored this advertising strategy.’ This dilemma needs to be managed carefully.

**Building Trust – Candor About Accuracy**

Honestly portraying the accuracy levels of your forecasts over time should build trust. Sometimes it does, sometimes it doesn’t. “The problem is that some managers don’t seem to acknowledge the inherent unpredictability of the world,” said Goodwin. “I’ve actually come across managers who expect forecasts to be 100 percent accurate. To them, showing any accuracy level less than that may detract from trust.”

Forecasters who live in the real world know that acceptable performance is context-dependent. For example, if you’re forecasting something for which you have masses of past data that follows regular seasonal patterns – such as the number of passengers flying from one airport to another – you might expect a 5 percent error rate or less. But if you’re trying to predict the stock market, which is highly random, you would expect a high degree of inaccuracy. The same is true of a new product launch, where you have very little historical data to go on, perhaps only data from analogous products launched at an earlier time. A 200 percent error rate would not be out of the question.

“It becomes a matter of educating managers about what level of accuracy can be expected in that environment,” said Goodwin. “One way of trying to improve this understanding is to show forecast accuracy against a benchmark method, for other products or the industry norm. These comparisons might give a sense of how much randomness and unpredictability is in what you’re trying to forecast – and show that your method outperforms the alternatives.”

**Building Trust – Candor About Uncertainty**

The Met Office – the UK’s national weather service – forecast that April 2012 would be a drier-than-average month, yet it turned out to be the wettest month on record. Newspaper headlines were unkind to the weather forecasters, but as Goodwin points out, the forecasters weren’t entirely wrong. Their internal forecasts were probabilistic – showing that there was a 25 percent chance the month would be dry, and less than a 15 percent chance the month would be very wet. The office had simply publicized the most likely event, without acknowledging publicly that their forecasts involved a lot of uncertainty.

Goodwin showed several ways forecasters can describe the uncertainty associated with forecast outcomes.

“**As human beings, we’re brilliant at explaining random events after the event, explaining why sales went up or down. But if you can find genuine explanations why there are variations from your forecast, it probably is worth it to explain to managers that there is no way that event could have been reasonably anticipated.”**

Paul Goodwin
Professor of Management Science, University of Bath

“**You might be using the best possible forecasting method, which is giving errors of 10 percent, but management might say, ‘I’m paying you to produce forecasts; why are your forecasts wrong? Most people think the world is more predictable than it really is. Some managers need to be educated that random events do occur that just can’t be predicted.”**

Paul Goodwin
Professor of Management Science, University of Bath
• **Prediction intervals** give a range of possible values and state how confident you are that the actual result will fall within the range. For example, we might say we are 90 percent sure that demand will be between 200 and 280 units.

• **Fan diagrams** take prediction intervals to another level, showing how confident you are that the actual result will be within a given range, shown at several confidence levels. For instance, a line chart forecast of inflation rates might show that there’s a 50 percent chance that inflation will be within one range, a 75 percent chance that it will fall in a (broader) range band, and a 90 percent chance that it will fall within the broadest range band.

![This example shows 50%, 75% and 90% ranges;](image)

**The short-term inflation forecast as a fan diagram**
Harmonized index of consumer prices
Change on year earlier, %

![Figure 1: A fan diagram shows the chance that the outcome will fall into a given range. Source: Bank of Finland Bulletin 2003](image)

• **Density forecasts** provide the most clarity about uncertainty, expressing the forecast as a probability distribution. For instance, you might see that potential sales can range from 15 to 65 units, with a probability attached to each figure. You might see that there’s little chance of sales being below 30 or above 50. Business users might prefer more specificity, but this visualization honestly acknowledges the uncertainty associated with possible outcomes.
Figure 2: A density forecast presents the possible outcomes as a probability distribution.

Building Trust – Competence

“People are not very good at assessing the competence of forecasters,” Goodwin said. “We very often believe expert forecasts without any real evidence that they are accurate, and we tend to remember the occasional very good or very bad forecasts, rather than general accuracy. The reputation of an exemplary forecaster may be destroyed by one isolated bad forecast that people readily recall. In contrast, one surprisingly accurate (and lucky) forecast of a major event that no one else foresaw will probably promote a poor forecaster to the status of a seer. There is actually evidence that those who do make these incredibly prescient forecasts of the next big event, the people that get these rare events right, actually tend to do worse on average than other forecasters.”

Establishing competence in the eyes of forecast users is therefore a matter of showing your long-term levels of forecast accuracy against benchmarks. It’s also a matter of being able to explain the rationale for choosing one forecasting method over another.

Trustworthy software helps. “The quality of forecasting software is extremely variable,” said Goodwin. “I even know of some commercial software that can’t display a graphical series. And of course, errors in the software products will lead to erroneous forecasts. Even the world’s most widely used forecasting product, Microsoft Excel, contains many errors in its statistical routines. … So choose your product carefully. A product that might look outwardly very pleasing may actually be giving you the wrong forecast.”
Balancing the Drivers of Trust

Goodwill, honesty, transparency and competence – these sound like indisputably good qualities in a forecaster. “The problem is that these drivers of trust may operate against each other,” said Goodwin. For example:

• If you’re honest that you can’t reduce forecast error below a certain level, some people may misinterpret the error rate as a sign of incompetence.
• Highly competent people may be drawn to advanced methods that, by being complicated, reduce transparency.
• Being transparent about uncertainty can reduce the user’s perception of confidence. “What, you’re a forecaster and you don’t know?” Some senior managers have unrealistic expectations and can’t understand why a forecast would have uncertainty at all. And they want direct answers. They don’t want to hear there’s a 90 percent chance sales will be between 300 and 348 units; they want to hear that sales next month will be 324 units.

“These difficulties mean that engendering trust requires political, communication and relationship skills in addition to technical skills,” said Goodwin. “Perhaps a slightly less accurate forecast that is accepted by users is better than a more accurate forecast that is simply thrown into the trash can. Sometimes it may be better to compromise slightly on accuracy than to push a highly accurate forecast that everyone refuses to use.”

Best Practice Recommendations

Benchmark Against Other Methods to Show Relative Accuracy

If 100 percent accuracy isn’t the goal – and it isn’t – forecasters would do well to compare their forecasts against those produced by other benchmark methods, to show that their forecasts are at least as accurate as can be expected.

At minimum, you can compare your forecast to the naïve forecast, which simply extrapolates forward from past data. For example, the naïve forecast says that if we sold 500 color televisions this month, we will sell 500 color televisions next month. “If you’ve not beating the naïve forecast, you’ve got real problems,” said Goodwin, while acknowledging that the naïve forecast can occasionally be the winning method for difficult-to-forecast environments, such as the stock market or currency exchange.

Wouldn’t it consume too much time to create multiple forecasts by different methods for each business question? Not necessarily, said Goodwin. “With modern software and computing power, you can produce alternative forecasts pretty quickly. It may be worth the extra effort to make sure your forecasts are trusted and used. I’d rather put in double the effort into getting my forecasts used, than put in half the effort and see my forecasts thrown into the trash can.”
Test the Robustness of the Forecasting Methodology

How do you know you’re using the right forecasting model for the business question at hand? First, ask whether the model makes sense – intuitively it seems reasonable – and then test it, said Goodwin. “It’s very easy to have a complex model that perfectly fits almost all of your past observations, so the fit of the forecast to past data isn’t a good guide. Leave some of your data hidden away from your model – out-of-sample or holdout sample data. Use your early data to fit your model, and then use your holdout sample to see how well the model forecasts with the data it didn’t see.”

You can also test the robustness of your forecasting model when applied to a wide range of data sets, different products or to a moving window of time.

Average the Results of Multiple Forecasting Methods

“One of the major findings of forecasting research over the last 20 to 30 years has been that taking an average of the forecasts generated by several methods tends to improve accuracy,” said Goodwin. He cited a powerful example on the Forecasting Principles website (forecastingprinciples.com), where multiple methods were averaged to predict the share of the vote in the last three US presidential elections. Methods ranged from the judgment of experts to opinion poll forecasts to econometric methods and more – which, when averaged, led to accuracy within 0.5 or 1 percent.

There are several reasons this hybrid approach works so well, said Goodwin, but it is probably because the different methods draw on different information sources. For example, a judgment-based forecast draws on the forecaster’s expertise and recent knowledge. The econometric model could draw on relationships with economic conditions, advertising and other factors. Averaging multiple methods also tends to cancel out biases. Where one method tends to forecast too high, and the other tends to forecast too low, the average of those two forecasts will be more accurate.

“You don’t normally need a sophisticated way of getting the combined forecasts,” said Goodwin. “Forecasting research has repeatedly shown that a simple average or mean of those forecasts often does well.”

Draw on the Talents of Both Quantitative and Business Specialists

Who produces better forecasts: statisticians who are skilled with the numbers, or business users with a deep knowledge of their products and customers? There are limitations to both, so they probably complement each other, said Goodwin. “The business user may suffer from motivational bias or optimism bias – ‘I believe in this product’ – or from cognitive biases that place too much weight on most recent observations. On the flip side, the mathematician locked away in a room, buried away in the numbers, lacks the depth of knowledge about the market and products, and won’t be able to forecast exceptional events.

Goodwin recommends averaging the forecasts produced by both sides – or having the mathematician forecast normal situations and the business expert step in when an exceptional situation such as a special event or promotion is on the horizon.”
Closing Thoughts

"An untrusted forecast is a worthless forecast, no matter how accurate it is," said Goodwin. "The creation of trust involves working to increase the forecast user’s belief and confidence in the reliability and integrity of your forecast. That involves affirming their perception of your goodwill. It involves sharing information in language they are comfortable with and reassuring them of your confidence in your systems and methods – while conveying the necessary degree of humility by acknowledging that no forecaster ever gets it ‘right’ every time.

"It also involves being transparent about your methodologies, providing clear and cogent explanations of your forecasts, being honest about uncertainty, and letting users review an honest history of your accuracy levels relative to reasonable benchmarks. If you do all this, chances are you will have the satisfaction of seeing your forecasts used to drive decisions, and you’ll avoid the demoralization that comes when your forecasts are adjusted or ignored."

About the Presenter

Paul Goodwin is a Professor of Management Science at the University of Bath, UK. He is a former director of the International Institute of Forecasters, editor of the International Journal of Forecasting, and a member of the editorial board of the Journal of Behavioral Decision Making. In addition to extensive research and publication, Goodwin has advised companies and government departments on forecasting and decision making. He also writes a column in Foresight aimed at practitioners.

For More Information

Read the Foresight article “Why Should I Trust Your Forecasts?” (M. Sinan Gönül, Dilek Öñkal and Paul Goodwin): foresight.org/foresight/forms/trust

View the SAS on-demand forecasting workshop, a virtual tour of SAS® Forecast Server: sas.com/reg/gen/corp/2188540

Learn more about SAS forecasting and econometrics software: sas.com/technologies/analytics/forecasting/index.html

Download free SAS white papers on forecasting: sas.com/technologies/analytics/forecasting/index.html#section=6

Read Foresight, the International Journal of Applied Forecasting: foresight.org/foresight
About SAS

SAS is the leader in business analytics software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions, SAS helps customers at more than 65,000 sites improve performance and deliver value by making better decisions faster. Since 1976 SAS has been giving customers around the world THE POWER TO KNOW®.