What is a Good Forecast?
A good forecast is one that has value to its user. Easy to say, but often hard to turn into an objective metric for evaluating forecaster performance. Business and investment decisions depend on many considerations apart from the forecast, so it’s frequently a challenge to identify the contribution of forecasts to the bottom line.

By default, most forecasters are judged on the basis of their average accuracy as most commonly measured by the mean absolute error (MAE), mean absolute percentage error (MAPE), or root mean square error (RMSE) of past forecasts. All give larger weight to larger errors and assume that users care equally about over- and under-predictions. But are large errors really more costly than small errors? Is excessive pessimism really as costly as excessive optimism? It is impossible to generalise. Everything depends on context, and specifically on the actions taken in response to the forecasts.

The Disconnect Between Accuracy and Profitability in Financial Markets
In financial markets, the value of a forecast can often be inferred directly from the amount of money that can be made by the action of trading on the forecast, possibly adjusted by the risk of these trades. It’s natural to ask, then: Are accurate forecasts also profitable forecasts? Against all intuition, the folklore in financial markets is that profitability is not related to RMSE accuracy, and not even to directional accuracy.

Why and when does this disconnect between accuracy and profitability arise? Let’s start by looking at the relation between accuracy and profitability.
and profitability within a group of financial analysts who made regular forecasts of interest rates in the 1990s.

Some time back, I collected the track records of 25 U.S. financial analysts. Most worked for banks or were advisors to fund managers and trading operations. They are the guys (yes, they’re all guys) that you see on CNN or Bloomberg News trying to second-guess the Fed. Every month for over a decade these analysts had made short- and long-term forecasts of the key 3-month Treasury bill yield. So it was easy, if tedious, to work out how accurate each of the forecasters had been. I used several error metrics, including the RMSE of their T-bill forecasts and their directional accuracy (DA), which is the percentage of times that the forecaster had correctly called the direction of the next move in interest rates.

There were three surprises.

First, the average directional accuracy of the forecasters was 48.5%, slightly worse than the toss of a coin – something to remember when you next see a talking head on some financial news channel.

Second, even the best forecasters were worse than a no-change benchmark forecast. Using the most recently published value of the T-bill yield and forecasting no change over the whole period covered by our data would have yielded a lower RMSE than any of the professionals in the panel.

Third, surviving for over 10 years within the group of analysts were some astonishingly bad forecasters, measured both on RMSE and DA criteria. Either forecasters have such high entertainment value that it obliterates memories of their past failures or our accuracy measurements are missing something.

The audience for these analysts consisted of investors and traders, so it is natural to try to measure the usefulness of the forecasts by seeing whether the forecasts could easily be used to make money in the T-bill market. As a first shot at this, imagine a trader buying and selling T-bill futures. The price of these futures is market-determined and is expressed as a number like 95, meaning that the market-implied future interest rate is 100-95 = 5%. Suppose that an analyst’s forecast is 4%. A trader who trusted this forecast would buy T-bill futures, expecting the price to rise from 95 (100-5%) to 96 (100 – 4%). Conversely, if the analyst’s forecast was 6%, the trader would sell T-bill futures, expecting the price to fall to 100-6 = 94.

Using this simple trading strategy over a 10-year period and calculating the profits made from following each analyst led to a couple of further surprises.

First, all of the analysts – even the very worst – generated positive profits.
Second, the profits made from following individual analysts were uncorrelated with root mean square error accuracy.

Figure 1 shows accuracy (RMSE) on the horizontal axis and profitability on the vertical axis. The forecasters’ names have been replaced by letter codes to protect the guilty, like forecasters XX and L who produced forecasts of interest rates with RMSEs above 2.5% (this at a time when typical annual changes were around 1.5%). Looking along the horizontal axis, we can see that the most accurate forecast is the no-change (SPOT+2) benchmark, that forecasters D and U were not far behind, and that the average (Consensus) forecast is also pretty accurate.

Looking at the vertical axis, which measures profits from trading, we can see that following any of the forecasters – even the least accurate – would have made positive profits. However, the very diffuse scatter of points tells us that the most accurate forecasters are not necessarily the most profitable. The no-change forecast is one of the least profitable forecasts. The Consensus forecast is only averagely profitable. The most profitable forecaster is Y, who is somewhere in the middle of the accuracy rankings. The correlation between profitability and accuracy is clearly pretty weak, and if we disregard the outliers XX and L, is close to zero.

**Little Misses and Big Hits**

The reasons for this disconnect become more apparent when we look closer at the behaviour of the T-bill yield and the detail of the forecasters’ track records.

The size of changes in T-bill yields varies over time. For maybe 90 percent of months, T-bill yields move up or down by only small amounts. The other 10 percent of the time – usually when the Fed has taken decisive action to loosen or tighten monetary policy – T-bill yields move by much larger amounts. During the months when T-bill yields had low volatility, our forecasters had trouble anticipating events and got the direction of change right only about 45% of the time. However, ahead of the really big moves, the forecasters almost all understood that a major policy change was imminent and got the direction of change right in around 80% of these months.

So their directional accuracy was indeed 45% x 90 + 80% x 10 = 48.5%, rather below what would be expected by chance. But their profits depended not so much on getting the direction of change right on average, but on getting it right ahead of the big market moves. In the 90% of months when changes in yields were small, the net losses made from trading on the forecasters’ inaccurate forecasts were small. In the months when major changes...
in yield occurred, the gains made from the forecasters' correct predictions were much larger, and more than offset the losses in other months.

In Figure 2, I have plotted a smoothed curve describing the distribution of monthly changes in T-bill yields between 1982 and 2008. The x-axis measures how many standard deviations (SD) observations are from their mean value of just below zero. Overlaid on the distribution of yields is the familiar normal distribution, which underpins so much statistical theory. Relative to the normal curve, we can see that there are many more very small changes than the normal curve would predict, and also many more extreme rises and falls of more than 2 or 3 standard deviations. Technically, we say that the T-bill distribution is “fat-tailed” and displays “excess kurtosis” (peakedness). Most financial time series – exchange rates, share prices, commodity prices – share this property, as do many economic time series. Growth rates are generally positive and not too variable. But every decade or so we experience a sharp recession or a short-lived bubble.

The significance of this for forecasting is that if changes in the target variable are normally distributed, we can expect a significant negative correlation between forecaster profitability and forecaster error (as measured by RMSE or MAE or MAPE); that is, higher error correlates with lower profitability. If forecasts and targets have approximately the same variability, the correlation of profits and RMSE will be about -0.5 (see the Technical Note at the end of this article).

So, generally, we could count on accurate forecasters also giving us valuable trading signals. However, if changes in the target variable are non-normal, the correlation vanishes because profits depend mainly on the correct prediction of the extreme changes in the target.

When I tracked the analysts' forecasts over time, it turned out that many of their errors happened because they correctly forecast that a major change in the target variable was coming but could not guess the timing. So all through the year ahead of a major interest rate hike, the forecasters would be predicting the rise in rates. During most of the year the actual T-bill yield was moving around randomly, so the forecasters looked pretty foolish and certainly no better than a no-change forecast. Eventually, when the rate rise happened, they were vindicated, but they had piled up a year's worth of mistakes on their track record in the meantime. In this sense, these guys were right at the right time.

**Lessons for the Mainstream Business Forecaster**

In the real world, many target variables (such as Sales) behave like these financial variables. Much of the time, variations in the target are small and unpredictable, but every so often a very large change occurs for an understandable reason.

However, the pay-off to sales forecasters of acting on forecasts of the large changes is very different from the payoffs to an analyst trading a futures price. Take a look at Figure 3, which shows monthly deviations from trend in a target variable through 2007 and 2008. Let's say that, at the start of 2007, forecasters correctly predict that there will be some favourable event during the year, but don’t know exactly when it will occur. Imagine that they simply forecast +10, say, every month until they are proved right (shown as the solid black line on the figure). In 2008, the forecasters suspect there will be some adverse event, and forecast -10 every month until the bad event occurs (the dashed line on the figure).
Note that these forecasts will be biased: in most months of 2007, too optimistic; in most months of 2008, too pessimistic. This phenomenon is known as the “peso problem.” During the late 1970s and 1980s, in spite of raging inflation, Mexico tried to maintain a fixed exchange rate against the U.S. dollar. Consequently, forecasters knew that each month there was a large probability of no change in the exchange rate and a small probability of a really massive devaluation. The “expected value” of the change in the exchange rate was for a small devaluation, so forecasters wound up looking biased towards pessimism in the many months when Mexico managed to keep its exchange rate pegged, and then too unadventurous in the few months when massive devaluations occurred.

You can dine out on an accurate forecast, but don’t expect it to pay for the dinner.

The blue line in Figure 3 shows the actual outcomes. These follow a fat-tailed distribution. In most months the changes are small, but in December 2007 (as forecast) there is a large positive movement, and in October 2008 (again as forecast) a large negative movement.

Now let’s compare the experiences of the trader and the sales forecaster. In March 2007, both were forecasting +10, and the outcome was -12. If the target were a “price,” the trader would have made a loss of 12 from betting on a price rise. If the target is a sales figure, the loss is harder to calculate, but likely to be larger, since it will depend not on the difference between the outcome and zero, but on the difference between the outcome and the forecast, in this case -12 - 10 = -22. For example, if sales forecasts are used to manage inventories, this -22 would represent the excess inventory resulting from the optimistic sales forecast. In April 2007 the trader makes some money, since although the price does not rise by 10, it does at least go up, and this will offset the loss made in March. The sales forecaster, though, is probably still not happy, since although the demand in April has been high, the business is still carrying excess inventory. By December 2007, the trader is very happy since the profit made from being on the right side of the market in December easily outweighs the (small) net losses through the year. The sales forecaster can perhaps feel vindicated but may still get criticised because now inventory may be too low.

The same argument applies in 2008. Because profits to the trader do not depend on the size of the forecast error but only on the correctness of the directional forecast, the trader will make small net losses through the year and a lot of money in October when the crash happens. The sales forecaster now may have insufficient inventory, and the benefits from anticipating the large fall in sales in October may or may not be enough to offset the costs of missing out on sales earlier in the year.

Even in the financial markets, companies may be unforgiving of the losses that tend to precede correct forecasts of large changes. The late Tony Dye was head of investments at leading UK fund manager Phillips and Drew (P&D) through the 1990s. From 1996 on, in spite of steadily rising stock markets worldwide, he repeatedly pointed out that, relative to expected company earnings, these markets were hugely overvalued and advised his clients to avoid equities. As their investments
with P&D underperformed year after year compared to competing funds, investors withdrew more and more money from the firm and eventually, in February 2000, Dye was fired. Days later, the bubble burst, and P&D shot from the bottom to the top of the fund performance league.

Like our analysts, Dye made a series of bad directional calls followed by a very correct call of the stock market correction. Unlike our analysts, the bad directional calls were all wrong in the same direction, so during the bubble Dye was racking up losses that were barely offset by the firm’s outperformance in the crash. Since directional accuracy and profitability were both pointing in the same direction, P&D can hardly be blamed for inferring that Dye was misunderstanding the dynamics of the stock market.

**Conclusion**

If you can put a monetary value on forecasts, then this monetary gain or loss should be used instead of, or at least alongside, arithmetical measures of forecast accuracy. When both indicators point in the same direction over a significant number of forecasts, then we can get a clear signal of the quality of the forecast. However, as we have seen, in theory and practice this need not be the case. Correlations between accuracy and profitability can be low. The most valuable forecasts are not always the most accurate. In these cases, profits and losses matter more than root mean square errors. You can dine out on an accurate forecast, but don’t expect it to pay for the dinner.

**Technical Note: Correlation of Error and Profitability**

Proof of this is a little complicated, but if the target follows a random walk with normally distributed changes, the correlation between RMSE and profitability from buying and selling fixed amounts of the underlying asset on directional signals is 

\[-\frac{2}{\sqrt{\pi}} \frac{s^2}{s^2 + \tau^2},\]

where \(\tau^2\) is the variance of the target and \(s^2\) the variance of the forecast. If \(s^2 = \tau^2\), this correlation is -0.56. If, more plausibly, forecasts are less variable than the target, the correlation will be weaker, but always negative.

**Roy Batchelor** is Foresight’s Financial Forecasting Editor and Professor of Banking and Finance at the Cass Business School of the City University of London. He has contributed some of this journal’s most provocative articles on forecast modeling, such as his discussion of “Worst-Case Scenarios in Forecasting” (Summer 2010 issue) and “Forecasting Sharp Changes” (Spring 2009).

r.a.batchelor@city.ac.uk
Putting collaboration into practice.

Join your peers, leading S&OP and supply chain practitioners, and noted researchers for 1.5 days of practical professional development. You’ll be immersed in real-world experiences and lessons learned about how you can access the benefits of internal collaboration that S&OP affords, as well as expanding collaborative endeavors with supply chain partners.

What can you expect from the Foresight Practitioner Conference?

Information and experiences that have been tested and proven in the real world. You’ll learn from successful practitioners whose insights come from firsthand experience. You’ll benefit from presentations of practical research from top business school professors, and you’ll find yourself thinking outside the box in an interactive, business school-style simulation.

Practical information you can put to use right away. You’ll return to your organization with a deeper understanding of the methods and practices that will equip you to enable better collaboration internally and across the supply chain. You won’t find vendor presentations or sales pitches in the program.

Opportunities to interact and network. Registration is limited to 100 to enable you to interact with our expert speakers and to ensure that you have maximum opportunities to network with your peers from over 50 companies working through challenges similar to yours.