Forecasting, causality and feedback

Rob J Hyndman

30 September 2022

Shmueli & Tafti (2022) are to be congratulated for producing a ground-breaking paper, bringing together causality and prediction in an innovative analysis, and casting some types of forecasts in a new light.

Understanding causality is important when forecasting

Forecasters have often considered causality as irrelevant, thinking that all we need is correlation in order to build predictive models, and whether that correlation is a result of causation or not does not matter much. This widespread view in the community perhaps explains the lack of forecasting papers that interact with the causal literature. Shmueli & Tafti (2022) have clearly shown that ignoring causality leads to misleading analysis in the case where forecasters can affect the outcome they are trying to forecast. I think it is also worth considering causality when building models, because causal variables are usually much better predictors than non-causal but correlated predictors.

To take a trivial example from Hyndman & Athanasopoulos (2021, Section 7.8), it is possible to forecast if it will rain in the afternoon by observing the number of cyclists on the road in the morning. When there are fewer cyclists than usual, it is more likely to rain later in the day. The model can give reasonable forecasts, not because cyclists prevent rain, but because people are more likely to cycle when the published weather forecast is for a dry day. In this case, there is a causal relationship, but in the opposite direction to our forecasting model: the number of cyclists falls because there is rain forecast. A better forecasting model for rainfall will not include cyclists, but it will include atmospheric observations from the previous few days, which of course are causal for subsequent rain. This simple example demonstrates that causality is an important consideration in building good forecasting models.

Forecasting with feedback

The particular framework discussed by Shmueli & Tafti (2022) involves the organizations that produce forecasts independently having the power to modify user behaviour. A related, but more common situation, is where the *forecasts* (rather than the *forecasters*) affect the outcome being forecasted — that is, the forecasts involve feedback (Hyndman & Athanasopoulos, 2021, Section 1.1). This is alluded to in the authors' last paragraph. Let's consider three examples where there is forecasting with feedback, to illustrate some of the complexities involved.

First, imagine a financial investment company publishing forecasts of certain stocks. Those forecasts can directly affect the prices of those stocks, moving them closer to the forecasts, if investors assume that the published forecasts represent an accurate reflection of value. In this case, the stock investment company may appear prescient, and attract new business, as a result of the accuracy of the forecasts. This is slightly different from the problem discussed by Shmueli & Tafti (2022), in that there is no direct action inducing the behaviour modification other than making the forecasts and publicising them. Nevertheless, we can still consider the act of publishing the forecast as an intervention intended to affect the outcome, and construct the causal diagram shown in Figure 1.

Published forecasts can also affect other forecasts, resulting in a herd-effect, where many forecasters follow the lead of some prominent forecasters, unwilling to be seen to be out-of-step (Bewley & Fiebig, 2002). When published forecasts can affect the outcome, herd behaviour only amplifies the effect, as it seems everyone is confidently in agreement with a specific outcome.

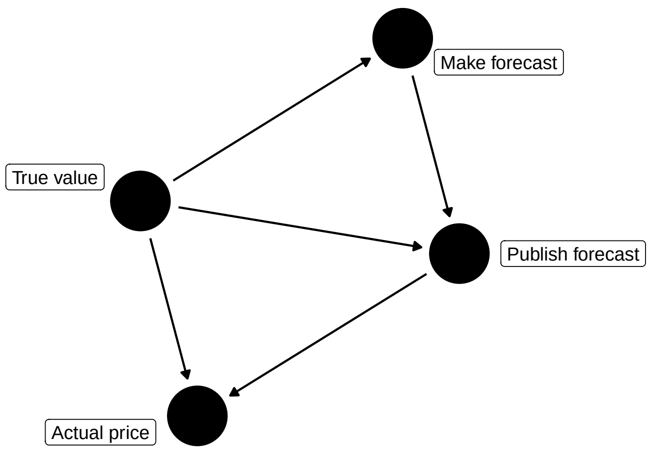

Figure 1: *Causal diagram representing the effect of publishing a forecast on a stock price.*

When accuracy is not the objective

Now think about an auction for fine art where the auction house releases a forecasted sale price for each item (usually specified as a range). Figure 1 also applies here: the published forecast can affect what someone is willing to pay, on the assumption that it is an accurate reflection of value.

However, unlike the stock price example, here the most desirable outcome is to increase the sale price, rather than minimize the expected prediction error. In fact, the published forecast may be different from the actual forecast, in order to have the desired effect on the sale price. Thus, while the problem appears similar in some ways to the social media setting and the stock price setting, in fact it is quite different from a forecasting perspective. Finding a good forecasting model is only of secondary interest, with the main objective being to find how to push the sale price as high as possible through manipulating the published forecast, and how it is presented to the prospective buyers.

Here, and in many other contexts, the objective of publishing a forecast is to affect the outcome, and in that case a successful forecast may be inaccurate.

COVID-19 forecasting

To take another example, forecasts of disasters are often intended to prompt a policy response to avoid the predicted outcome. Every week since March 2020, I have been involved in producing forecasts of daily COVID-19 cases for all Australian state and territory governments (Moss et al., 2022). These forecasts are provided to the senior health officers of each state and territory, who advise their governments of appropriate policy responses. A forecast of a high level of COVID-19 cases has sometimes resulted in the government imposing tighter restrictions or even a lockdown, and the resulting numbers of COVID-19 cases is much lower than we forecast. Here, our forecasts are made subject to existing policies continuing into the forecast period, and a change in policy means the forecast model assumptions are no longer true. This is a successful outcome, not a forecasting failure.

Our forecasts have largely not been made public, but are presented confidentially to relevant government authorities. Our objective function is to minimize forecast error (measured using CRPS) subject to existing health policies continuing. However, when framed in the causal modelling framework, it is conceivable that a better objective function might be to minimize total cost to the community (measured as some combination of public health and economic costs), and then the published forecasts may be quite different from the forecasts that minimize CRPS. Of course, we don't do this — I am simply pointing out how the causal framework leads naturally to considering behavioural manipulation beyond minimizing forecast accuracy.

Evaluating forecasts with feedback

When forecasts involve behaviour modification or feedback, we need to take special care when evaluating the resulting forecasts. Shmueli & Tafti (2022) have provided some helpful analysis of how to approach this problem when behaviour modification is intended to reduce forecast error. When behaviour modification has some other objective, usually resulting in an increased forecast error, a different approach is needed.

One possibility is to only evaluate outcomes that were realized under the same conditions that existed at the time of forecasting (without any behaviour modification or indirect effects due to forecast publication). However, this is only possible if there is available data that have no behaviour modification effects.

Another approach is to include the environmental conditions as covariates in the model, essentially shifting the behaviour modification $do(B)$ into the exogenous variables $\boldsymbol{x}$. Then the results can be evaluated ex post based on the conditions that existed at the time of the observed outcome. However, that is only possible if there is historical data available to estimate the effect of the change-of-policy intervention when training the model.

A third possibility is to use counterfactuals, where a model is built ignoring the behaviour modification, and the forecasts are then judgementally adjusted based on the estimated effect of the intervention. This is similar to what we did in Athanasopoulos, Hyndman & O'Hara-Wild (2022) in forecasting Australian tourism numbers in a post-pandemic world. Again, this is only possible when there is data available without behaviour modification effects.

For COVID-19 forecasts, a mixture of the first two of these approaches has been used. For forecasts up to about seven days ahead, it is reasonable to assume that any policy responses have negligible effect given the incubation period and time frame for disease development, and forecasts can be evaluated as usual assuming the conditions haven't changed since the forecasts were produced. For longer-term forecasts, we can compare the accuracy of ex post forecasts (which take account of covariates describing the potential policy responses) with ex ante forecasts (assuming no change in the environment). The difference provides information about the adequacy of the models as well as the effectiveness of the policy response.

Thanks

Thanks to Galit Shmueli and Ali Tafti for producing such an interesting and stimulating paper. I would also like to acknowledge their assistance in commenting on an early draft of this commentary, helping me clarify my thinking.

References

Athanasopoulos, G, RJ Hyndman & M O'Hara-Wild (2022). Probablistic forecasts using expert judgement: the road to recovery from COVID-19. *J Travel Research*. to appear.

Bewley, R & DG Fiebig (2002). On the herding instinct of interest rate forecasters. *Empirical Economics* **27**(3), 403–425.

Hyndman, RJ & G Athanasopoulos (2021). *Forecasting: principles and practice*. 3rd ed. Melbourne, Australia: OTexts. `OTexts.org/fpp3`.

Moss, R, DJ Price, N Golding, P Dawson, J McVernon, RJ Hyndman, FM Shearer & JM McCaw (2022). Forecasting COVID-19 activity in Australia to support pandemic response: May to October 2020. *medRxiv*. doi.org/10.1101/2022.08.04.22278391.

Shmueli, G & A Tafti (2022). How to "improve" prediction using behavior modification. *International Journal of Forecasting*. to appear.