International Journal of Forecasting (In Press: 2015)

Order effects in judgmental forecasting

Zoe Theocharis and Nigel Harvey

University College London

Address for correspondence:
Nigel Harvey
Department of Cognitive, Perceptual and Brain Sciences
University College London
Gower Street
London WC1E 6BT
UK
Tel: +44 207 679 5387
Fax: +44 207 436 4276
Email: n.harvey@ucl.ac.uk

Running head: Order effects

Order effects in judgmental forecasting

Running head: Order effects

Abstract

In two experiments, forecasters made a sequence of five forecasts from different types of time series, either from the nearest horizon to the most distant one (1, 2, 3, 4, 5) or in one of two other orders, both of which required the forecast for the most distant horizon to be made first ('end-anchoring'). These latter two orders differed in terms of the direction of the remaining forecasts: it was a horizon-increasing order (1, 2, 3, 4) or a horizon-decreasing one (4, 3, 2, 1). End-anchoring improved forecast accuracy, especially for more distant horizons. It resulted in the trajectory of the forecast sequence being closer to the optimal one. Direction of forecasting after end-anchoring affected forecast quality only when the optimal trajectory of the forecast sequence displayed strong nonlinearity. End-anchoring provides a simple means of enhancing judgmental forecasts when predictions for a number of horizons are made from each series.

Key words: judgmental forecasting, time series, forecasting practice, evaluating forecasts, forecasting education

There is now a large corpus of research on judgmental forecasting (Lawrence, Goodwin, O'Connor and Önkal, 2006). Although some studies have found judgmental forecasts to be as accurate as statistical ones (e.g., Lawrence, Edmundson and O'Connor, 1985), others have shown them to be subject to a number of systematic errors or biases (e.g., Carbone and Gorr, 1985; Sanders, 1992). Examples of such errors include those occurring when people damp trends by producing a sequence of forecasts with a trend that is less steep than that in the data series (Eggleton, 1982; Harvey and Reimers, 2013; Lawrence and Makridakis, 1989), when they overestimate the degree of sequential dependence in independent series (Bolger and Harvey, 1993; Eggleton, 1982; Reimers and Harvey, 2011), when they add noise to their sequence of forecasts in proportion to the level of noise in the data series (Harvey, 1995; Harvey, Ewart and West, 1997), and when they make higher forecasts for desirable than for undesirable outcomes, possibly because action can be expected to be taken to prevent increases in the latter case (Eggleton, 1982; Lawrence and Makridakis, 1989; Harvey and Reimers, 2013).

Techniques that have been found to improve judgmental forecasts include training with feedback (Goodwin and Fildes, 1999; Goodwin, Önkal-Atay, Thomson, Pollock and Macaulay, 2004), decomposition of the forecasting task by making separate estimates for each component of the time series and then combining them (Edmundson, 1990; Webby. O'Connor and Edmundson, 2005), combining predictions from a number of forecasters (Clemen, 1989), and taking advice (Lim and O'Connor, 1995; Goodwin, Gönül and Önkal, 2013). Unfortunately, most techniques of this sort impose heavy additional costs in terms of time, money or effort.

The above findings are based primarily on research into point forecasts for the immediate horizon (i.e. the next data point after the most recently observed one). There have been fewer studies into judgmental forecasts for more distant horizons (i.e. data points beyond the most recently observed one). Here, we have two main aims. First, we shall investigate factors that differentially affect accuracy of forecasts for different horizons. Second, we shall use our findings to identify a very low cost technique for improving judgmental forecasts, particularly those for more distant horizons.

In the next section, we provide a brief review of research on horizon effects on judgmental forecasting. Then we introduce the two factors that we manipulate in our experiments and outline our hypotheses about their effects.

Effects of forecast horizon

Uncertainty increases as we progress into the future. Hence, except when data series possess unique short-term volatility characteristics (Thomson, Pollock, Henriksen and Macauley, 2004), both statistically based forecasts and judgmental forecasts tend to be worse for more distant forecast horizons (Lawrence et al, 1985). Rate of deterioration, measured by increase in mean absolute percentage error (MAPE), is broadly similar for the two types of forecast (Lawrence, Edmundson and O'Connor, 1986) but reasons for it differ. As we saw above, judgmental forecasts, unlike most statistical forecasts, show trend damping. This causes their directional error to increase over the forecast horizon (Harvey and Reimers, 2013; Lawrence and Makridakis, 1989). What potential explanations are there for this increase in error with increasing horizon?

To make forecasts for the first horizon, people appear to use the last data point as a mental anchor and then make some adjustment away from that point to take account of the pattern in the series. Typically, these adjustments are insufficient[1]. As a result, trend damping is observed with trended series and forecasts from non-trended series appear to exaggerate the sequential dependence in the data. Furthermore, people add random noise to the result of the anchoring and adjustment process to produce their forecasts (Harvey, 1995; Harvey et al, 1997). One possible explanation for this phenomenon is that forecasters add noise, either intuitively or deliberately, to make their sequence of forecasts look more representative of the presented data series (Harvey, 1995).

Bolger and Harvey's (1993) analysis of their experiments showed that forecasts for later horizons are made in a similar way except that the previous forecast rather than the last data point is used as a mental anchor. If their account is correct, then the random noise added to previous forecasts would accumulate as people make forecasts for increasingly distant horizons. If such an accumulation of random noise could be eliminated, forecasts for these more distant horizons would improve in accuracy and variability across forecasters in the trajectory of the forecast sequence would be reduced.

End-anchoring

The theoretical analysis presented in the previous section suggests that forecasting performance would be improved by preventing forecasts for horizons beyond the first one being made in sequence and, thereby, accumulating the random errors associated with each one. One obvious way of doing this is to ask forecasters to make their forecast for the most distant horizon first. We might assume that forecasters do this by using the anchoring and adjustment heuristic that is normally used to make an initial forecast. For example, for trended series, instead of making a forecast for the first horizon by anchoring on the last data point and adjusting away from that value by a proportion (P) of the difference between the last two data points (Bolger and Harvey, 1993), they could make a forecast for, say, the fifth horizon by anchoring on the last data point and adjusting away from that value by 5P (i.e. five times the size of the adjustment used when forecasting for the first rather than the fifth horizon).

Forecasters may find making an initial forecast for the most distant horizon more difficult than making an initial forecast for the first horizon and it may take them a little longer. However, once that forecast has been made, their task is transformed from one of extrapolation to one of interpolation.

We would expect this manipulation to produce its greatest improvement on the forecast for the most distant horizon. This is the horizon that would be most affected by accumulation of noise components in previous forecasts. However, because interpolation is a more constrained task than extrapolation, the end-anchoring produced by making an initial forecast for the most distant horizon may also improve forecasts for less distant horizons. To make the intervening forecasts, people may simply use linear interpolation between the last data point and their forecast for the most distant horizon. We would still expect them to add a noise component to the results of each forecast in this interpolation (Harvey, 1995) but this would not determine the trajectory of the forecast sequence.

Based on the above rationale, we will test the following hypotheses.

H_1: Requiring forecasters to make their initial forecast for the most distant horizon will produce more accurate forecasts for that horizon than when they make their forecast for it last.

H_2: Requiring forecasters to make their forecast for the most distant horizon first rather than last will also increase the accuracy of forecasts for less distant horizons.

Reversing the direction of the forecasting

Once forecasters have made their initial forecast for the most distant horizon, they could proceed in one of two ways. They could forecast from the end of the data series towards their existing forecast for the most distant horizon. So forecasts for five horizons would be made in the order: 5, 1, 2, 3, 4, where lower numbers represent horizons closer to the end of the data series. We shall refer to this as *horizon-increasing forecasting*. Alternatively, they could make forecasts in the reverse direction, working from their initial forecast for the most distant horizon back towards the end of the data series. Thus, when forecasts for five horizons were required, they would make them in the order: 5, 4, 3, 2, 1, where lower numbers again represent horizons closer to the end of the data series. We shall refer to this as *horizon-decreasing forecasting*. There are reasons to suppose that direction of forecasting will influence accuracy but that the effect of this variable will depend on the characteristics of the underlying signal in the time series and on the level of noise overlaid on that signal.

First, consider forecasting from series containing linear trends. Trend damping effects tend to be greater with downward than with upward trends (Harvey and Bolger, 1996; Lawrence and Makridakis, 1989; O'Connor, Remus and Griggs, 1997). Hence, if we assume that an upward trend when forecasting in a horizon increasing manner (i.e. left to right) is transformed to a downward trend when forecasting in a horizon decreasing manner (i.e. right to left), errors in forecasting upward trends should be larger when people forecast in a horizon-decreasing manner than when they forecast in a horizon increasing manner.

Second, suppose that the final point of an autocorrelated data series has been perturbed well away from the mean or trend line of the series by noise. When forecasting in a horizon increasing manner, forecasters could take the effects of autocorrelation into account (Reimers and Harvey, 2011): for example, if the last point of an untrended series with a first order autocorrelation of .5 was 8 points above the series mean, optimal forecasts for the next three horizons would be four, two, and one point above the mean. However, when forecasting in a horizon decreasing manner, people may be unable to make any allowance for autocorrelation.

Third, with untrended independent data series, there is no obvious reason to expect any major asymmetries between horizon increasing and horizon decreasing modes of forecasting if interpolation is reasonably good. However, if it is poor (perhaps because people have some difficulty taking into account the position of the anchor they are moving towards), forecast errors for horizon 1 may be larger for horizon decreasing than for horizon increasing forecasting whereas errors for horizon 4 may be larger for horizon increasing than for horizon decreasing forecasting. More generally, we expect the order in which interpolation forecasts are made to affect accuracy because of noise introduction and accumulation effects. Asymmetries of the sort described above are likely to be present when forecasting from any series containing an underlying signal overlaid by noise.

These three suggestions are examples of how forecasting direction may influence accuracy. However, there may be other as yet unidentified factors related to misperception of the signal or to failure to exclude noise that differentially affect forecasts made in a horizon increasing and in a horizon decreasing manner. Therefore, the hypotheses that we test are not highly specific.

H_3: Forecasting accuracy in horizon increasing conditions will be different from that in horizon decreasing ones.

H_4: The effects of a reversal in the direction in which forecasts are made will depend on the characteristics of the data series.

Experiment 1

In this experiment, participants were presented with time series comprising 35 points and asked to make forecasts for the next five points. To test the above hypotheses, we manipulated the horizon for which the initial forecast was made (first versus last), the direction of forecasting when the forecast for the final horizon was made first (horizon increasing versus horizon decreasing), and series type.

Method

Participants One hundred and twenty students (48 men, 72 women) from University College London (UCL) acted as participants. They were recruited from UCL's pool of participants. We provided a brief description of the experiment and members of the pool registered to take part in the study. Their mean age was 26 years and they had all passed courses in statistics (thereby rendering our findings more relevant to professional settings). They were paid a flat rate of £1.00 for their participation[2].

Design Participants were randomly allocated into two groups. The first group (no end-anchoring) made their forecasts for the five horizons in the order in which the data points appeared (i.e. 1, 2, 3, 4, 5). The second group (end-anchoring) did not. Instead they made their forecast for the most distant horizon (i.e. 5) first. In this second group, there were two sub-groups. The horizon increasing forecasting sub-group made their forecasts from the end of the data series towards the forecast that they had already made for the most distant horizon. Thus, their five forecasts were made in the order 5, 1, 2, 3, 4. In contrast, the horizon decreasing forecasting sub-group made their forecasts in the reverse direction moving from their initial forecast for the most distant horizon back towards the final point of the data series. Thus, their forecasts were made in the order 5, 4, 3, 2, 1.

All participants made predictions for four different types of time series. Hence, they each produced a total of 20 forecasts (five horizons for each of four types of series). Characteristics of the four types of series are described in the next section.

Stimulus materials The four types of series were: an untrended series of independent data points; an untrended series of highly autocorrelated data points; a series of independent data points with a linear trend imposed upon them: a series of independent data points with a seasonal signal imposed upon them. Series were presented graphically. Examples of the four types of series can be seen in Figure 1. Each panel in the figure shows 35 data points (seen by participants) followed by five optimal forecasts (not seen by participants).

These types of series represent most of the basic components from which more complex series are constructed. We wanted to identify the effects of our manipulations on these basic constituents. Real data series, with complex underlying structures, would have been have been less appropriate for this purpose. We would expect that findings related to a particular type of series described by one of these basic patterns (e.g., seasonal series) would generalize to more complex series containing that type of pattern as a constituent.

Figure 1 about here

Untrended series were constructed by inserting appropriate parameters into the following generating equation: $X_t = \alpha X_{t-1} + (1 - \alpha)\mu + \varepsilon_t$, where X_{t-1} was the previous observation, μ was the mean of the series, α was the degree of autocorrelation ($\alpha = 0.9$ for autocorrelated series and $\alpha = 0$ for random series), and ε was noise produced by randomly drawing values from a Gaussian distribution with a mean of zero and a variance of σ^2 ($\sigma^2 = 30.0$ for both autocorrelated and independent series). The mean value, μ, was selected to ensure that the final data point was close to the vertical mid-point of the screen.

Linear trended series were produced by using the equation: $X_t = 5t + \varepsilon_t$. Its noise term, ε, had a mean of zero and a variance of 19.0. The final data point of these trended series was approximately 10% of the screen height above its vertical mid-point.

Seasonal series were constructed by using the equation: $X_t = 70\cos(100t) + 170 + \varepsilon_t$, where the noise term had a mean of zero and a variance of 225. The starting point of these series was chosen so that the last data point was a) close to the vertical mid-point of the screen and b) one third of the way from the mid-point of the seasonal cycle towards its peak (Figure 1). Each wavelength phase lasted for 12 time periods. There were 3.25 wavelengths in the screen. Each wavelength's width corresponded to a 30% of the screen width.

In this experiment where attention is primarily focused on the effects of task characteristics with series with different signals, noise levels were fairly low but still sufficient to ensure that participants had uncertain percept of the signal. Seasonal series contained higher noise levels than those imposed on other types of series because of the larger amplitude of the series' underlying signal.

The scale used for data presentation might affect graphical perception (Legge, Gu and Luebker (1989) and so we used the optimality suggestions for graphical displays identified by Cleveland, McGill and McGill (1988) to guide our choice (Figure 1). This ensured that there was sufficient vertical space available to make forecasts, especially with trended series (c.f. Lawrence and

Makridakis, 1989). The phase of the seasonal series was selected to enable us determine whether participants detected the seasonal pattern (Figure 1).

Time series were generated uniquely for each participant and the four types of series were randomly ordered separately for each of them. The task was not performed within a particular scenario, such as one associated with sales forecasting, to avoid introduction of frame-specific biases (e.g., elevation biases arising from optimism or perceived control effects). Hence, the vertical axes of the graphs used to present the series were unlabelled.

Series were presented as line graphs. After the end of each series, five vertical lines were presented in the next five time periods to indicate where forecasts had to be made. When a forecast was made by clicking on one of the vertical lines, a blue dot appeared in the position of the cursor when the mouse was clicked.

Procedure Each participant performed the task individually on a computer in a separate cubicle. They read a short introduction to the study and then entered their demographic details (age, sex). They were instructed to view each series and then click on each of the vertical lines to show where they expected future points in the series to appear.

The no end-anchoring group received the following instructions:

“HELLO AND WELCOME TO THE FORECASTING SKILLS EXPERIMENT”

You will be presented with a set of time-series, which represent consecutive values of a variable. The y axis represents the values of the series and x axis represents time in days. Each time a time-series appears, you will be asked to indicate your predictions (5 predictions in each trial) by clicking on the vertical lines after the last data point. So, after observing the given series, which contains values for 35 consecutive days, you will have to forecast what will happen in the next 5 days (Day 36, 37, 38, 39 and 40). Blue points will appear where you click. After indicating all forecasts for the next five days, your forecasts will be stored and the next time series will appear. You will be presented with four series in all.”

Participants in the end-anchoring conditions received similar instructions but they were also informed of the specific order in which they had to make their forecasts. Furthermore, the task was constrained to ensure that their forecasts were actually made in that order. Thus the programme would accept a forecast (and show a blue dot to indicate its position) only if it was made in the required order. Participants were not allowed to change their forecasts or go back to previous screens.

For participants in the no end-anchoring group, all five vertical lines were presented at the same time as the series. Forecasts were made from the nearest horizon to the most distant one in the order 1, 2, 3, 4, 5. As forecasts were made, a blue line linked each new forecast with the last data point (forecast for horizon 1) or with the immediately preceding forecast (all other forecasts).

For participants in the horizon decreasing sub-group of the end-anchoring group, all five vertical lines were again presented at the same time as the series. However, an explicit screen message prompted participants to make their forecasts in a horizon decreasing manner (in the order 5, 4, 3, 2, 1). As forecasts were made, a blue line linked each new forecast with its predecessor.

Participants in the horizon increasing sub-group of the end-anchoring group first made their forecast for the most distant horizon for each of the four series. Thus, initially, only a single vertical line at furthest horizon was presented with each series to signal that only the forecast for that horizon was required. Once that forecast had been made for all series, participants returned to each one (presented in the same order as before) to make the remaining four required forecasts working from the end of the data series towards the furthest horizon for which they had already made their forecast. To enable them to do this, the remaining four vertical lines appeared on the screen at this point to indicate the positions of these required forecasts[3]. Thus forecasts were made in the order 5, 1, 2, 3, 4. As forecasts were made, a blue line linked each new forecast with the last data point (forecast for horizon 1), with the previous forecast (forecasts for horizons 2 and 3), or with both the immediately preceding forecast and the forecast for the most distant horizon (forecast for horizon 4).

Results

All participants succeeded in completing the task. However, we excluded participants whose forecasts were at least 3 inter-quartile ranges from the median of each group. This resulted in a total of 116 participants, 58 in each of the two conditions.

To test H_1 and H_2, we compare mean absolute error (MAE) between Group 1 (no end-anchoring) and Group 2 (end-anchoring). MAE is the mean across participants of the absolute difference between the forecast and the actual value of the series, which, in turn, was generated by the above equations. To cast more light on the effects of end-anchoring, we also report supplementary analyses a) of the Mean Directional Error (MDE) and b) of regressions of forecasts for each type of series on to horizon for each participant. Then, to test H_3 and H_4, we compare MAE between Group 2a (horizon increasing forecasting after end anchoring) and Group 2b (horizon decreasing forecasting after end-anchoring). We also report the same type of supplementary analyses as for end-anchoring.

Effects of end-anchoring Graphs of MAE in the two conditions are shown in Figure 2 for each of the four series types. They show accuracy decreasing with increasing horizon and the decrease appears to be higher in the no end-anchoring group for seasonal, linear trended, and autoregressive series. To examine the significance of these effects, we carried out separate two-way analyses of variance (ANOVA) on the MAE data for each series type, using horizon as a within-participants variable (five levels) and condition (end-anchoring, no end-anchoring) as a between-participants variable. Here and later, Huynh-Feldt corrections were applied to address violations of sphericity. Where the effect of horizon or the interaction between that variable and condition was significant, we also tested for the simple effects of condition at each horizon.

Figure 2 and Table 1 about here

Results of these analyses are shown in the upper part of Table 1. We can see that effects of horizon are significant for all series types[4] and that, for seasonal, linear and autocorrelated series, there were significant effects of condition and of the interaction between that variable and horizon. The tests of simple effects indicate that end-anchoring was particularly effective for later horizons (4 and 5).

For all series that contain a pattern as well as noise, these analyses are consistent with our first hypothesis (H_1) that end-anchoring improves the accuracy of the forecast for the most distant horizon: in each case, the simple effect of group was significant for horizon 5. The results are also consistent with our second hypothesis (H_2) that end-anchoring also improves accuracy of forecasts for less distant horizons: for series containing patterns, end-anchoring facilitated forecasting for either or both horizons 2 and 4.

We now report two supplementary analyses designed to throw light on the reasons for these effects. The first analysis is of Mean Directional Errors (MDE) for each series type (Figure 3). Directional errors are calculated as actual forecast minus optimal forecast[5]. Hence, the increasing directional error for forecasting the downward section of the seasonal series and the decreasing directional error for forecasting the upward sloping linear trended series are both evidence of trend damping. It is immediately apparent from Figure 3 that one effect of end-anchoring is to reduce trend damping.

Figure 3 and Table 2 about here

Two-way ANOVAs on MDE data for each series type, using horizon as a within-participants variable (five levels) and condition (end-anchoring versus no end-anchoring) confirmed that this was so. The upper part of Table 1 shows effects of condition only on those series containing linear or seasonal trends where damping effects could impair forecast accuracy. In these analyses, the effects of horizon indicate trend damping and the interaction demonstrates that end-anchoring reduces that effect.

The second supplementary analysis was designed to investigate the effect of end-anchoring on the slope of the sequence of five forecasts in more detail. We used an approach employed by Harvey and Reimers (2013). We fitted linear regression models to each one of the four sequences of five forecasts produced by each participant. Thus, for each sequence, we fitted the model: forecast = a + b (horizon) + error. Then, for each series type, we used t-tests to examine whether the constants (a) and trend coefficients (b) in the two conditions differed from one another and whether each of them differed from the optimal values derived from the generating equation. We also tested whether the variance of the coefficients and the error variance in the model were greater in the no end-anchoring group than in the end-anchoring group.

Values of coefficients in each condition and in the generating equation and levels of error variance in each condition are shown in Table 2 for each series type. This table also indicates the comparisons that reached significance. However, we will highlight the main results here.

For the seasonal series, the mean slope of the forecast sequence was significantly more negative (i.e. steeper) in the end-anchoring group than in the no end-anchoring group ($t(114) = 6.557$; $p < 0.05$). For the linearly trended series, the mean slope was significantly more positive (i.e. steeper) in the end-anchoring group than in the no end-anchoring group ($t(114) = -3.519$; $p < 0.001$). These results confirm that end-anchoring acts to decrease trend damping.

Variance of the trend coefficients was significantly lower in the end-anchoring group than in the no end-anchoring group for linearly trended (F (57, 57) = 1.78; $p < .05$) and autocorrelated series (F (57, 57) = 2.22; $p < .05$). (Data for the other two series types are in the same direction but the comparisons did not attain significance.) This shows that there was a tendency for end-anchoring to reduce the degree to which the slope of the forecast sequence drifted away from its correct value.

Effects of direction of forecasting

One sub-group made forecasts in a horizon increasing sequence after end-anchoring: horizons were forecast in the order 5, 1, 2, 3, 4. The second sub-group made forecasts in a horizon decreasing sequence after end-anchoring: horizons were forecast in the order 5, 4, 3, 2, 1. Here we test hypotheses H_3 and H_4 by comparing overall forecast error (MAE) in the horizon increasing and horizon decreasing sub-groups. Graphs of MAE in the two conditions are shown in Figure 4 for each of the four series types. We carried out separate two-way ANOVAs on each of them using horizon as a within-participants variable and condition as a between-participants variable.

Results of these analyses are shown in the upper part of Table 3. While all types of series showed significant effects of horizon, only seasonal series showed a significant interaction between that variable and forecast direction. Thus, results for seasonal series are consistent with our third hypothesis: effects of direction of forecasting affected accuracy for that series type. Furthermore, the results as a whole are consistent with our fourth hypothesis: effects of direction of forecasting depended on series type.

____

Table 3 and Figures 4 and 5 about here

____

Why were only seasonal series influenced by direction of forecasting? For the other three types of series, reasonably good forecasts could be made by making linear interpolations between the last data point and the forecast already made for the most distant horizon. This was because the sequence of outcomes that required forecasting was linear. In contrast, the sequence of outcomes that had to be forecast in the seasonal series was non-linear: its values first increased markedly and then decreased. However, forecast sequences did not show this pattern: the value of the first two forecasts stayed close to that of the last data point; values of later forecasts then decreased more slowly than the values of the outcomes to be forecast. In other words, the forecast sequence did not show such a clear point of inflection as the sequence of outcomes that had to be forecast: it was more linear than it should have been. These patterns are shown in Figure 5.

These impressions were confirmed in regression analyses. Using a step-up procedure, we found that 40 of the 58 participants' forecasts showed significant linear components. In those cases, the linear models explained an average of 86% of the variance. Adding a quadratic component significantly increased the variance explained by the model in only three of these 40 participants and, on average, it explained only an additional 10% of the variance. Also, comparing the 40 models (with quadratic components included) to a model of the outcomes to be forecast (produced by continuing the generating function) showed that the coefficient for the quadratic component was significantly lower in the participants' forecast sequences (t (39) = 6.10; $p < .001$) than in the sequence of outcomes. These analyses imply that the interpolations that participants made between the last data point and the forecast that they had already produced for the most distant horizon were more linear than they should have been.

As Figure 5 shows, the near-linearity of participants' sequence of forecasts reflected their failure to increase the value of their first few forecasts above the value of the last data point. As a result, their forecasts were too low – and the extent to which they were too low was greater in the horizon decreasing forecasting group than in the horizon increasing forecasting group. A two-way ANOVA on MDE confirmed this. It showed a significant effect of forecast horizon (F (2.45, 136.96) = 27.21; $p < .001$) and an analysis using polynomial contrasts indicated that it had linear and quadratic components. There was also a significant effect of forecasting direction (F (1, 56) = 16.94); $p < .001$)

and a marginally significant interaction between the two variables (F (2.45, 136.96) = 2.78; p = .055). Tests for simple effects showed that the effect of forecasting direction was significant at horizon 2 (F (1, 56) = 5.84; $p < .05$), horizon 3 (F (1, 56) = 7.90; $p < .01$), horizon 4 (F (1, 56) = 9.68; $p < .01$) and horizon 5 (F (1, 56) = 10.68; $p < .01$).

Why did this pattern of results occur? It appears that participants forecasting in a horizon decreasing manner anchored their judgments on the low value of the forecast that they had already made for the most distant horizon. As a result, although they then increased the value of their forecasts for horizons 4, 3, and 2, they did so insufficiently. Their forecast for horizon 1 was then made by linearly interpolating between their forecast for horizon 2 and the last data point. In contrast, those forecasting in a horizon increasing manner anchored on the relatively high value of the last data point. As a result, their forecasts for horizons 1, 2, and 3 were made at the same level as that point. Then, their forecast for horizon 4 was made by linearly interpolating between their forecast for horizon 3 and the forecast that they had made earlier for horizon 5.

Thus when our participants had to forecast a nonlinear sequence of four outcomes between the end of the data series and a forecast that they had earlier made for the most distant horizon, they used an initial strategy based on anchoring to make the first three of those forecasts and then made the final forecast by linear interpolation. This produced different levels of accuracy for horizon decreasing and horizon increasing forecasting.

Discussion

The experiment showed that, when the data series contain a pattern, judgmental forecasts for a sequence of outcomes can be improved by making the forecast for the most distant horizon first. It also showed that, when that is done, the order in which the remaining forecasts are made does not matter if the sequence of outcomes that require forecasting lie in a straight line. However, if they contain some other (i.e. nonlinear) pattern, accuracy of forecasts made in a horizon increasing direction (from the last data point towards the previously produced forecast for the furthest horizon) can differ from forecasts made in a horizon decreasing direction (from the previously produced forecast for the furthest horizon towards the last data point). We shall discuss these findings in turn.

End-anchoring Making the forecast for the furthest horizon first clearly improved accuracy not only of that forecast but of other forecasts too. This result was expected. We found that it occurred for two reasons. First, as our regression analyses showed, the trajectory of the forecast sequence became less variable. We expected this because, without end-anchoring, each forecast after that for the first horizon is based on its noisy predecessor but is not constrained by a forecast for a more distant horizon: the task is one of extrapolation. In contrast, end-anchoring constrains the forecast trajectory: the task is one of interpolation.

Second, as our analyses of MDE and our regressions show, end-anchoring reduced trend damping. This effect was present both in series with upward linear trends and in the downward sections of seasonally trended series[6]. We suspect that it occurred because participants found forecasting in the end-anchoring condition more difficult. As a result, they devoted more cognitive resources to the task and performed it better. To check this account, we compared the mean time taken to make the first forecast in the two groups. This analysis showed that it was less in the no end-anchoring group (4.37 seconds) than in the end-anchoring group (6.96 seconds) ($t (221.79) = 12.35$; $p < .001$). We also compared the time to make all five forecasts in the no end-anchoring group (9.60 seconds) with horizon decreasing forecasting sub-group of the end-anchoring group[7] (13.69 seconds) and found it to be greater in the latter case ($t (65.46) = 7.29$; $p < .001$). These two analyses confirm that forecasters devoted more cognitive resources to their task in the end-anchoring condition.

This finding can be interpreted in terms of models that posit different modes of cognitive processing: an intuitive system that acts rapidly, heuristically, non-consciously, and with little effort and a deliberative system that acts slowly, analytically, consciously, and with effort (Kahneman, 2011).

Thus forecasting in the normal way from the data series for increasingly distant horizons may be an intuitive process that relies on anchoring heuristics and produces ‘biases’ such as trend damping. In contrast, making forecasts for the most distant horizon first is likely to be a slower, more cognitively demanding, deliberative process that is less susceptible to the sort of biases produced by heuristic processing.

Direction of forecasting After end-anchoring, the forecasting task was transformed from one of extrapolation to one of interpolation. When linear interpolation was appropriate (random, autocorrelated or linearly trended series), there was no difference in accuracy between interpolating in a horizon increasing manner from the end of the data series towards the anchor provided by the forecast for the most distant horizon and interpolating in a horizon decreasing manner from that anchor towards the end of the data series. However, when linear interpolation was not appropriate (seasonal series), interpolating in a horizon decreasing manner produced higher levels of error than forecasting in a horizon increasing manner.

The reason for this appears to be that people adopted different strategies for forecasting in the two cases. Differences seem to be related to the direction of forecasting rather than to task differences (i.e. the horizon increasing condition was performed in two phases). We can see this if we pay closer attention to participants’ behaviour when forecasting from the seasonal series. The section of the seasonal series that had to be forecast comprised the peak of a cycle followed by a descending segment (Figure 5). When forecasting in a horizon decreasing manner, the descending segment became an ascending one and was forecast in the same way as a linear trend. For the first three forecasts they made (horizons 4, 3, and 2), participants anchored on the forecast that they had made immediately before and then adjusted upwards to take the trend into account. As there adjustments were insufficient, some trend damping was observed (Figure 5). Then they made their final forecast for horizon 1 by linearly interpolating between their previous forecast for horizon 2 and the last data point.

When forecasting in a horizon increasing manner, participants approximated the peak of the seasonal series as an untrended linear series and forecast it as if it were one. Thus, their forecasts for horizons 1, 2, and 3 were forecast at the same level as the last point of the data series. Then they made their final forecast for horizon 4 by linearly interpolating between the forecast that they had just made for horizon 3 and the forecast that they had made earlier for horizon 5. This strategy for forecasting was, unlike the one for horizon decreasing forecasting, not subject to trend damping: it therefore produced forecasts that were higher and closer to the target outcome series (Figure 5).

Experiment 2

In this experiment, we examine the effects of a) increasing the level of noise in the data series and b) changing the phase of the seasonal series so that the sequence of outcomes that had to be forecast was approximately linear rather than nonlinear.

Increasing noise in the data series is likely to impair forecasting performance. However, higher levels of noise in series produce greater trend damping effects (Eggleton, 1982; Harvey and Bolger, 1996). As a result, a manipulation that reduces trend damping should still improve accuracy more when series noise is higher. Hence, we test the following hypothesis.

H_5: Although higher levels of noise will depress forecast accuracy, effects of end-anchoring will still be evident.

Requiring people to forecast an approximately linear section of the seasonal series should eliminate the difference between horizon decreasing and horizon increasing forecasting sub-groups of the end-anchoring condition. This is because linear interpolation, forecasters' default strategy after end-anchoring, would be as appropriate as it is for linearly trended or untrended autocorrelated series. We would expect it to be used irrespective of forecasting direction. Hence the higher levels of MDE that we observed for horizon decreasing forecasting from seasonal series in Experiment 1 should no longer obtain.

H_6: Requiring participants to forecast a linear rather than a nonlinear sequence of outcomes will eliminate the effect of forecasting direction on MDE.

Method

Participants Participants comprised 120 students (57 men, 63 women) drawn from the same pool as before. Their mean age was 28 years and they were paid the same amount of money as before for their participation (£1.00).

Design As the end-anchoring effect did not occur when there was no pattern in the data, we excluded random series from this experiment. In all other respects, the design was identical to that outlined for Experiment 1.

____

Figure 6 about here

____

Stimulus materials For the seasonal trended series, the amplitude of the seasonal variation was doubled: the equation used to generate these series was therefore $X_t = 140\cos(100t) + 170 + \varepsilon_t$. Also the variance of the noise component was increased by a factor of four to 900. The starting point of these series was chosen so that the last data point was a) close to the vertical mid-point of the screen and b) at the peak of the seasonal cycle (Figure 6).

The linearly trended series and the autocorrelated series were generated in the same way as in Experiment 1, except that the variance of the noise was increased by four times to a value of 120 in the former case and to a value of 76 in the latter one.

Procedure The procedure was identical to that used in Experiment 1.

Results

Analysis of MAE showed that forecasting was worse in this experiment than in the previous one (Seasonal series: $F(1, 228) = 497.21$; $p < .001$; Trended series: $F(1, 228) = 64.22$; $p < .001$; Autocorrelated series: $F(1, 228) = 125.70$; $p < .001$). It also deteriorated more with increasing horizon (Seasonal series: $F(2.72, 619.76) = 128.03$; $p < 0.001$; Trended series: $F(3.17, 724.31) = 5.15$; $p < 0.001$; Autocorrelated series: $F(2.01, 458.31) = 25.02$; $p < 0.001$). Given that series in the current experiment contained more random noise, these effects were to be expected. However, according to $H_{5,}$ effects of end anchoring should be maintained despite this reduced accuracy. To test this hypothesis, we again compared MAE in the no end-anchoring and end-anchoring groups.

To test H_6, we compared MAE in the horizon increasing forecasting and horizon decreasing forecasting sub-groups of the end-anchoring group to determine whether the effects of forecast direction for seasonal series that were apparent in Experiment 1 were eliminated in the current experiment.

Effects of end-anchoring Graphs of MAE in the two conditions are shown in Figure 7 for each of the three series types. They show accuracy decreasing with increasing horizon and the decrease again appears to be higher in the no end-anchoring group for seasonal, linear trended, and autoregressive series. To examine the significance of these effects, we carried out separate two-way analyses of variance (ANOVA) on the MAE data for each series type. Results of these analyses are shown in the lower part of Table 1.

Figure 7 about here

Again, all series types showed significant effects of horizon. They also all showed a significant effect of end-anchoring and/or an interaction between that variable and horizon. Tests of simple effects again showed that end-anchoring always had an effect on the most distant horizon (5), a result consistent with H_1. They also showed that, for seasonal series, end-anchoring improved accuracy of forecasts for less distant horizons, a result consistent with H_2. Thus, despite deterioration of overall forecast accuracy, effects of end-anchoring were maintained (H_5).

To throw light on the reasons for these effects, we now report the same two supplementary analyses that we carried out for Experiment 1. The first analysis was carried out on MDE (again calculated as actual forecast minus optimal forecast) for each series type (Figure 8). The increasing directional error for forecasting the downward section of the seasonal series and the decreasing directional error for forecasting the upward sloping linear trended series are both evidence of trend damping. It is clear that end-anchoring again acted to reduce trend-damping in these series (Figure 8).

____

Figure 8 and Table 4 about here

____

Results of two-way ANOVAs on MDE, using horizon as a within-participant variable and end-anchoring condition as a between-participants one, are shown in the lower part of Table 1. Effects of horizon (indicating trend damping) were again present only in those series containing seasonal or linear trends. Effects of condition were present only at longer horizons, thereby indicating that end-anchoring produced a reduction in trend damping in these types of series.

To carry out the second analysis, we again fitted regression models to each one of the four sequences of five forecasts produced by each participant. As before, for each sequence, we fitted the model: forecast = a + b (horizon) + error. Mean values of constants, trend coefficients and residual variance in each condition, together with optimal values derived from the generating equations are shown in Table 4. Also shown is the significance of statistical comparisons between the two groups and between each of them and the values in the generating equations.

For seasonal series, there was evidence that trend damping was reduced in the end-anchoring condition. In other words, the mean absolute value of the linear trend coefficient (b) was significantly lower in participants' forecast sequences in both conditions than in the generating equation and also lower in forecast sequences in the no end-anchoring condition than in the end-anchoring condition (Table 4). The effect did not reach significance for the linearly trended series. Also, in this experiment, there was no statistical evidence that the variability of b coefficients was greater in the no end-anchoring condition though, for all three series types, the difference across conditions was numerically in that direction.

Effects of direction of forecasting Figure 9 shows MAE scores for each of the three series types. Results of separate two-way ANOVAs on each series type, using horizon as a within-participants variable and forecasting direction as a between-participants variable, are shown in the lower part of Table 3. They indicate significant effects of horizon for all series types.

Trend damping contributed to the effects of horizon on MAE for seasonal and linearly trended series (Figure 10). Thus, as shown in the lower part of Table 3, two-way ANOVAs showed effects of forecasting horizon on MDE in the seasonal and linearly trended series but not in the autocorrelated series. Effects of forecasting direction and the interactions between this variable and horizon did not reach significance for any series type.

____

Figures 9 and 10 about here

____

In Experiment 1, horizon decreasing forecasting of seasonal series produced higher MAE scores than horizon increasing forecasting. We showed that this occurred because participants' strategy for horizon decreasing forecasting of the highly nonlinear sequence of outcomes was different from their strategy for horizon increasing forecasting of that sequence. In particular, horizon decreasing forecasting produced lower forecasts and, hence, higher MDE scores. In contrast, this experiment showed no effect of forecasting direction on MDE for seasonal series. This was because participants' linear interpolation strategy for forecasting the near linear sequence of outcomes was appropriate and the same for both horizon increasing and horizon decreasing conditions. As a result, MDE scores were no higher when participants were forecasting in a horizon decreasing manner than when they were forecasting in a horizon increasing one. Thus, these results are consistent with H_6.

Discussion

As expected, increasing the noise in the data series impaired forecasting and this impairment was greater for more distant horizons. This additional noise also resulted in effects of end-anchoring being only marginally significant for linear series. However, the more powerful cross-experiment comparison showed that the effect of end-anchoring was either maintained (linearly trended and untrended autocorrelated series) or magnified (seasonal series). End-anchoring had its effect by reducing trend damping effects (just as it did in Experiment 1). These effects tend to be greater with noisier series (Harvey and Reimers, 2013) and, as a comparison of Tables 1 and 2 shows, the b coefficient in forecast sequences underestimated the coefficient in the continuation of the data series by a larger amount here than in Experiment 1. Hence, the greater effect of end anchoring on seasonal series in this experiment may be attributed to the fact that there was more trend damping to be reduced in this experiment.

Direction of forecasting after end-anchoring did not affect forecast accuracy. This is in accord with H_6. It implies that the original effect found in Experiment 1 arose because the section of the seasonal series that required forecasting was strongly non-linear. In this experiment where the section of the seasonal series that required forecasting was close to linear, no effect of direction of forecasting was obtained. In other words, the original effect was not caused by the type of series represented by the data (seasonal rather than linearly trended) but by the characteristics (linear or nonlinear) of the ideal forecast sequence. In seasonal series, these characteristics depend both on the phase of the seasonal cycle at which forecasting must start and on the length of the forecast sequence.

General discussion

Our primary aim was to investigate the effects of end-anchoring. We anticipated that it would lead to improvements in the accuracy of judgmental forecasts. We know that people add noise to their forecasts (Harvey, 1995) and, when making a sequence of forecasts in order from nearest to most distant horizon, they use their previous forecast as a mental anchor (Bolger and Harvey, 1993). As a result of these two phenomena, a sequence of forecasts may be akin to a random walk and drift away its original trajectory. By requiring the most distant horizon to be forecast first, our aim was to eliminate this drift. Our first experiment did indeed show that end-anchoring reduced the variability across participants of the trajectories of forecast sequences made from the same underlying pattern.

However, the end-anchoring manipulation also reduced trend damping. This may have been related to forecasters finding their task more difficult. Making an initial forecast for five periods ahead is more challenging than making an initial forecast for one period ahead. Kahneman (1973) has argued that people cope with increased difficulty by allocating more cognitive resources to their task; for example, they may switch from using a rapid, heuristic, non-conscious, intuitive mode of processing to a slower, more analytic, conscious, deliberative mode of processing (Kahneman, 2011). The latter approach, though slower, tends to be more accurate. We suggested that end-anchoring improves accuracy because it results in more cognitive resources being devoted to the forecasting task (perhaps via a change from intuitive to deliberative processing). In support of this account, we demonstrated that initial forecasts took over fifty percent longer to produce in the end-anchoring group than in the no end-anchoring group.

Our second experiment used noisier data series. Forecasts were worse, showed greater trend damping, and deteriorated more rapidly as the forecast horizon increased. However, end-anchoring still decreased trend damping and, therefore, increased forecast accuracy for more distant horizons. In this experiment, variability of the forecast trajectories across participants was not significantly reduced by end-anchoring. Noisier data series produce noisier forecasts, which, in turn, reduce the likelihood of real effects attaining significance.

The experiments had a secondary aim. This was to investigate the effects of the direction in which forecasts were made after end-anchoring. There were plausible reasons to expect such a manipulation to have an effect on forecast accuracy, though we recognised that the nature of any effect was likely to depend on series type. Thus, for different types of series, we compared the accuracy of forecasts made in the order 54321 with that of those made in the order 51234. In fact, our results showed that the effect of forecast direction depended not on the type of series from which forecasts were made but on whether the optimal sequence of forecasts was linear or nonlinear. Forecast direction had an effect on accuracy only when that sequence was strongly nonlinear. In this case, horizon decreasing forecasting from the end-anchor (54321) produced higher levels of error than horizon increasing forecasting towards the end anchor (51234). This result could be explained by assuming that participants produced their first three forecasts after the end anchor by using (imperfect) extrapolation and then produced their final forecast by linear interpolation.

Implications for practice

A wide variety of techniques have been developed for improving judgmental forecasts. They include feedback-based training (Goodwin and Fildes, 1999), decomposition (Edmundson, 1990), combining forecasts from a number of forecasters (Clemen, 1989), and use of advisors (Lim and O'Connor, 1995). However, all of these approaches require quite heavy investments of time, money, or effort. Here we have shown that significant gains in forecast accuracy can be achieved simply by changing the order in which forecasts are made. In particular, requiring the forecast for the most distant horizon to be made first is an effective way of increasing the accuracy of forecasts, especially those for more distant horizons.

There may also be other advantages to requiring the most distant horizon to be forecast first. We pointed out above that this approach tends to place greater cognitive demands on judgmental forecasters and so may result in them using a slower, more analytic and deliberative mode of processing (Kahneman, 2011). Such deeper processing typically results in greater understanding of material and better memory for it (Bjork and Bjork, 2011; Craik and Lockhart, 1972). Thus practitioners who use the approach may be better able to appreciate the characteristics of patterns in the time series and also better at retaining them. This, in turn, may allow them greater insight into why those patterns occur.

Limitations

Our recommendations are relevant only when forecasters produce at least four or five forecasts from each data series. Advantages in terms of accuracy generally increase as forecast horizon extends into the future; accuracy for close horizons is unaffected by changes in forecast order. Nevertheless, future research could usefully explore the limits of the phenomenon. Do advantages of end-anchoring continue to increase for horizons beyond five? Are there certain types of series for which its advantages are present even for close horizons?

There has to be some pattern in the data series for order of forecasting to influence accuracy. If forecasting well merely requires mental extraction of the mean of an untrended random series, there is no advantage to be gained by end-anchoring (Experiment 1).

Although costs of end-anchoring are low relative to other techniques for improving judgmental forecasting, the technique imposes a greater cognitive load on forecasters and increases the time that they require to make their forecasts by about fifty percent.

The experiments reported here were conducted using student participants who had received instruction in statistics, a factor that could be expected to render the results more relevant to professional settings. We know that practitioners may perform worse than student participants

(Leitner and Schmidt, 2006), better than them (Önkal, Yates, Simga-Mugan and Ötzin, 2003), or no different from them (Lawrence, Edmundson and O'Connor, 1985). However, if we are correct in interpreting our findings as indicative of fundamental cognitive mechanisms that decrease forecasting biases, we should expect them to generalise to professional populations, even if the absolute level of performance of practitioner forecasters is different from that of our participants.

Future research

It would be worth determining the extent to which our findings generalize to forecasting tasks that depend on different types of display or different task environments. For example, Harvey and Bolger's (1996) findings that trends are damped more when data are presented in a tabular than in a graphical format suggests that reductions in trend damping due to end-anchoring could have a greater effect with a tabular format. It would also be worth exploring whether the optimism biases that have been observed when people forecast desirable variables (e.g., Harvey and Reimers, 2013) are attenuated by end-anchoring.

Such investigations could usefully be extended to studies of judgmental adjustment of forecasts where similar optimism and contextual phenomena have been documented (Fildes, Goodwin, Lawrence and Nikopoulos, 2009). End-anchoring by making adjustments to the furthest horizon first may also improve performance in these tasks: indeed, Franses and Legerstee's (2011) finding that adjustments made for distal horizons are more accurate than those made for proximal ones suggest that it would. It is also possible that availability of statistical forecasts that properly account for seasonality would attenuate the effects documented in Figure 5.

End-anchoring has implications for the study of cognitive processes underlying forecasting (Harvey, 2007), for popular consensus perceptions of time series such as those involved in climate change (Lewandowski, 2011), and for momentum strategies in asset markets (Hong, Lim and Stein, 2000). Future research examining the importance of end-anchoring in these contexts would benefit from use of real time series as such series are likely to be less stable than simulated ones and forecasting from them may not show the same biases (Lawrence and O'Connor, 1995).

Footnotes

1. When trends are very shallow, the opposite of trend damping ('anti-damping') is observed (Harvey and Reimers, 2013).

2. Harvey and Reimers (2013) found that performance characteristics when forecasting from time series with no causal factors was unaffected by presence of performance-related pay, a finding consistent with Camerer and Hogarth's (1999) meta-analytic review of the effects of financial incentives in different types of task.

3. Providing a single vertical line in the initial phase and the remaining four vertical lines in the second phase of the experiment enhanced task comprehension in both phases: in the former case, participants were immediately aware of the distant forecast required of them without having to discriminate between or count vertical lines; in the latter case, they were able to understand that the remaining four vertical lines corresponded to the remaining forecasts that they had to make

4. For all types of series, separate one-way ANOVAs showed the effect of horizon to be significant in the no end-anchoring condition and in the end-anchoring condition.

5. This optimal forecast was derived from the algorithms use to generate time series and so is different from the best forecast that participants could have made because they were not fully aware of the data generation process.

6. Though we did not investigate whether comparable reductions in trend damping occur with downward trended linear series and with upward sections of seasonal series, we are confident that they would. In fact, as trend damping effects are typically greater with downward trended linear series (e.g., Harvey and Bolger, 1996), it is likely that the reduction in damping produced by end-anchoring would be larger with such series than we found it to be with upward trended linear series.

7. We excluded the horizon increasing sub-group of the end-anchoring condition from this comparison because the procedure used to elicit forecasts delayed their production (see above).

References

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56-64). New York: Worth Publishers.

Bolger, F. & Harvey, N. (1993). Context-sensitive heuristics in statistical reasoning. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 46,* 779-811.

Camerer, C. & Hogarth R. (1999). The effects of financial incentives in experiments: a review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19*, 7-42.

Carbone, R. & Gorr W.L. (1985). Accuracy of judgemental forecasting of time series. *Decision Sciences, 16,* 153-160.

Clemen, R.T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559-583.

Cleveland, W., McGill, M., & McGill, R. (1988). The shape parameter of a two variable graph. *Journal of the American Statistical Association, 83*, 289-300.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Edmundson, R.H., (1990). Decomposition: a strategy for judgmental forecasting. *Journal of Forecasting, 9,* 301-314.

Eggleton, I.R.C. (1982). Intuitive time-series extrapolation. *Journal of Accounting Research, 20,* 68-102.

Fildes R., Goodwin P., Lawrence M., & Nikolopoulos K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting, 25*, 3-23.

Franses, P. & Legerstee, R. (2011). Experts' adjustment to model-based SKU-level forecasts: Does the forecast horizon matter? *Journal of the Operational Research Society, 62*, 537-543.

Goodwin, P., Önkal-Atay, D., Thomson, M. E., Pollock, A. C., & Macaulay, A. (2004). Feedback-labelling synergies in judgmental stock price forecasting. *Decision Support Systems, 37*, 175– 186.

Goodwin, P. & Fildes, R. (1999). Judgmental forecasts of time series affected by special events: Does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making, 12*, 37-53.

Goodwin, P., M.S. Gönül, and D. Önkal. (2013). Antecedents and effects of trust in forecasting advice. *International Journal of Forecasting,* 29, 354-366.

Harvey, N. (1995). Why are judgments less consistent in less predictable task situations? *Organizational Behavior and Human Decision Processes, 63,* 247-263.

Harvey, N. (2007). Use of heuristics: Insights from forecasting research. *Thinking and Reasoning, 13,* 5-24.

Harvey, N. & Bolger, F. (1996). Graphs versus tables: Effects of data presentation format on judgmental forecasting. *International Journal of Forecasting, 12,* 119-137.

Harvey, N., Ewart, T. & West, R. (1997). Effects of data noise on statistical judgment. *Thinking and Reasoning, 3*, 111-132.

Harvey, N. & Reimers, S. (2013). Trend damping: Under-adjustment, experimental artifact, or adaptation to features of the natural environment? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 589-607.

Hong, H., Lim T. & Stein J. (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance, 55*, 265-295.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. (2011). *Thinking fast and slow.* New York, NY: Ferrar, Straus and Giroux.

Lawrence, M & O'Connor, M. (1995). The anchor and adjustment heuristic in time-series forecasting. *Journal of Forecasting*, *14*, 443-451.

Lawrence, M., Edmundson, R & O'Connor, M. (1985). An examination of the accuracy of judgmental extrapolation of time series. *International Journal of Forecasting, 1,* 25-35.

Lawrence, M., Edmundson, R. & O'Connor, M. (1986). The accuracy of combining judgmental and statistical forecasts. *Management Science, 32*, 1521-1532.

Lawrence, M., Goodwin, P., O'Connor, M. & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting, 22*, 493-518.

Lawrence, M. & Makridakis, S. (1989). Factors affecting judgmental forecasts and confidence intervals. *Organizational Behavior and Human Decision Processes, 43*, 172-187.

Legge, G. E., Gu, Y., & Luebker A., (1989). Efficiency of graphical perception. *Perception & Psychophysics, 46,* 365-374.

Leitner, J., & Schmidt, R. (2006). A systematic comparison of professional exchange rate forecasts with the judgemental forecasts of novices. *Central European Journal of Operations Research, 14*, 87-102.

Lewandowsky, S. (2011). Popular consensus: Climate change is set to continue. *Psychological Science*, *22*, 460-463.

Lim, J. & O'Connor, M. (1996). Judgmental forecasting with interactive forecasting support systems. *Decision Support Systems, 16,* 339-357.

O'Connor, M., Remus, W., & Griggs, K. (1997). Going up - going down: How good are people at forecasting trends and changes in trends? *Journal of Forecasting, 16,* 165-176.

Önkal, D., Yates, J.F., Simga-Mugan, C. & Ötzin, Ș. (2003). Professional vs. amateur judgment accuracy: The case of foreign exchange rates. *Organizational Behavior and Human Decision Processes, 91*, 169-185.

Reimers, S. & Harvey, N. (2011). Sensitivity to autocorrelation in judgmental time series forecasting. *International Journal of Forecasting, 27*, 1196-1214.

Sanders, N. R. (1992). Accuracy of judgmental forecasts: A comparison. *Omega International Journal of Management Science, 20,* 353-364.

Thomson, M. E., Pollock, A. C., Henriksen, K. B., & Macaulay, A. (2004). The influence of forecast horizon on judgemental probability forecasts of exchange rate movements. *European Journal of Finance, 10*, 290-307.

Webby, W., O'Connor, M., & Edmundson, B. (2005). Forecasting support systems for the incorporation of event information: An empirical investigation. *International Journal of Forecasting, 21,* 411 – 423.

Table 1. Summary of effects of end-anchoring and horizon for each type of series in Experiments 1 and 2.

Series Type	Error Type	End-anchoring	Horizon	Interaction
Experiment 1				
Seasonal	MAE	*	** *(2, 5)*	**
	MDE	**	**	**
Linear	MAE	*	** *(2, 4, 5)*	*
	MDE	NS	**	**
Autocorrelated	MAE	*	** *(4, 5)*	**
	MDE	NS	NS	NS
Random	MAE	NS	**	NS
	MDE	NS	NS	NS
Experiment 2				
Seasonal	MAE	*	** *(3, 4, 5)*	*
	MDE	*	** *(4, 5)*	*
Linear	MAE	§	** *(5)*	NS
	MDE	*	** *(4, 5)*	NS
Autocorrelated	MAE	NS	** *(5)*	*
	MDE	NS	NS	NS

* Significant effects, $p < .05$

** Significant effects, $p < .001$

§ Significant effects, $p < .075$

Note: Italicized numbers in the Horizon column indicate horizons for which tests of simple effects showed effects of end-anchoring.

Table 2. Linear regressions of forecast sequences for each series type: mean values (variances in parentheses) of constants, trend coefficients and residual error variances. Actual values in the generating equations are shown for comparison.

		Constant (a)	Trend (b)	Error (e)
Seasonal	Actual	270.86	-20.50	
	No end-anchoring	228.69**† (23.48)	-2.14**† (9.71)	149.86
	End-anchoring	240.03**† (19.30)	-13.57**† (9.05)	235.12
Linear	Actual	207	5	
	No end-anchoring	211.45*† (8.60)	2.97*† (3.51) †	14.31
	End-anchoring	206.45† (7.22)	4.94† (2.63) †	16.25
Autocorrelated	Actual	150	0	
	No end-anchoring	148.51 (9.78)	0.86 (7.51) †	31.55
	End-anchoring	150.32 (15.02)	0.65 (5.03) †	44.88
Random	Actual	150	0	
	No end-anchoring	144.96 (25.52)	1.24 (7.74)	311.03
	End-anchoring	151.75 (17.25)	0.42 (6.36)	275.73

*Mean value different from that in the generating equation, $p < .05$

**Mean value different from that in the generating equation, $p < .01$

† Values differ between the no end-anchoring and end-anchoring groups, $p < .05$

Table 3. Summary of effects of direction and horizon for each type of series in Experiments 1 and 2.

Series Type	Error Type	Direction	Horizon	Interaction
Experiment 1				
Seasonal	MAE	NS	** *(2, 3, 4)*	*
	MDE	**	** *(2, 3, 4, 5)*	§
Linear	MAE	NS	**	NS
	MDE	NA	NA	NA
Autocorrelated	MAE	NS	**	NS
	MDE	NA	NA	NA
Random	MAE	NS	**	NS
	MDE	NA	NA	NA
Experiment 2				
Seasonal	MAE	NS	**	NS
	MDE	NS	**	NS
Linear	MAE	NS	**	NS
	MDE	NA	**	NS
Autocorrelated	MAE	NS	**	NS
	MDE	NS	NS	NS

* Significant effects, $p < .05$

** Significant effects, $p < .001$

§ Significant effects, $p < .075$

Note: Italicized numbers in the Horizon column indicate horizons for which tests of simple effects showed effects of end-anchoring.

Table 4. Linear regressions of forecast sequences for each series type: mean values (variances in parentheses) of constants, trend coefficients and residual error variances. Actual values in the generating equations are shown for comparison.

		Constant (a)	Trend (b)	Error (e)
Seasonal	Actual	322.14	-60.49	
	No end-anchoring	305.71**† (57.67)	-11.38**† (22.81)	1443.71
	End-anchoring	313.41**† (58.77)	-22.56**† (22.06)	1155.03
Linear	Actual	207	5	
	No end-anchoring	210.04 (12.76)	1.42** (4.49)	125.11
	End-anchoring	210.20* (11.79)	2.68** (4.16)	95.74
Autocorrelated	Actual	150	0	
	No end-anchoring	149.68 (20.83)	0.34 (15.88)	144.16
	End-anchoring	151.02 (35.75)	-0.46 (13.71)	260.35

*Mean value different from that in the generating equation, $p < .05$

**Mean value different from that in the generating equation, $p < .01$

† Values differ between the no end-anchoring and end-anchoring groups, $p < .05$

Figure captions

Figure 1. Experiment 1: Examples of the four types of series, showing 35 data points (seen by participants) followed by five optimal forecasts (not seen by participants) for seasonally trended, linearly trended, autocorrelated, and random series (clockwise from top left).

Figure 2. Experiment 1: Graphs of mean values of absolute error (together with standard error bars) in the no end-anchoring group (continuous lines) and in the end-anchoring group (dashed lines) for seasonal, linearly trended, autocorrelated, and random series (clockwise from top left).

Figure 3. Experiment 1: Graphs of mean values of directional error (together with standard error bars) in the no end-anchoring group (continuous lines) and in the end-anchoring group (dashed lines) for seasonal, linearly trended, autocorrelated, and random series (clockwise from top left).

Figure 4. Experiment 1: Graphs of mean values of absolute error (together with standard error bars) in the horizon increasing forecasting sub-group (continuous lines) and the horizon decreasing forecasting sub-group (dashed lines) for seasonal, linearly trended, autocorrelated, and random series (clockwise from top left).

Figure 5. Experiment 1: Graphs showing optimal forecasts (continuous lines) and participants' mean forecasts (together with standard error bars) in the horizon increasing forecasting sub-group (dashed lines) and the horizon decreasing forecasting sub-group (dotted lines) for seasonal, linearly trended, autocorrelated, and random series (clockwise from top left).

Figure 6. Experiment 2: Examples of the three types of series, showing 35 data points (seen by participants) followed by five optimal forecasts (not seen by participants) for seasonal (top panel), linearly trended (middle panel) and autocorrelated series (lower panel).

Figure 7. Experiment 2: Graphs of mean values of absolute error (together with standard error bars) in the no end-anchoring group (continuous lines) and in the end-anchoring group (dashed lines) for seasonal (top panel), linearly trended (middle panel) and autocorrelated series (lower panel).

Figure 8. Experiment 2: Graphs of mean values of directional error (together with standard error bars) in the no end-anchoring group (continuous lines) and in the end-anchoring group (dashed lines) for seasonal (top panel), linearly trended (middle panel) and autocorrelated series (lower panel).

Figure 9. Experiment 2: Graphs of mean values of absolute error (together with standard error bars) in the horizon increasing forecasting sub-group (continuous lines) and the horizon decreasing forecasting sub-group (dashed lines) for seasonal (top panel), linearly trended (middle panel)and autocorrelated series (lower panel).

Figure 10. Experiment 2: Graphs showing optimal forecasts (continuous lines) and participants' mean forecasts (together with standard error bars) in the horizon increasing forecasting sub-group (dashed lines) and the horizon decreasing forecasting sub-group (dotted lines) for seasonal (top panel), linearly trended (middle panel) and autocorrelated series (lower panel).

Figure 1

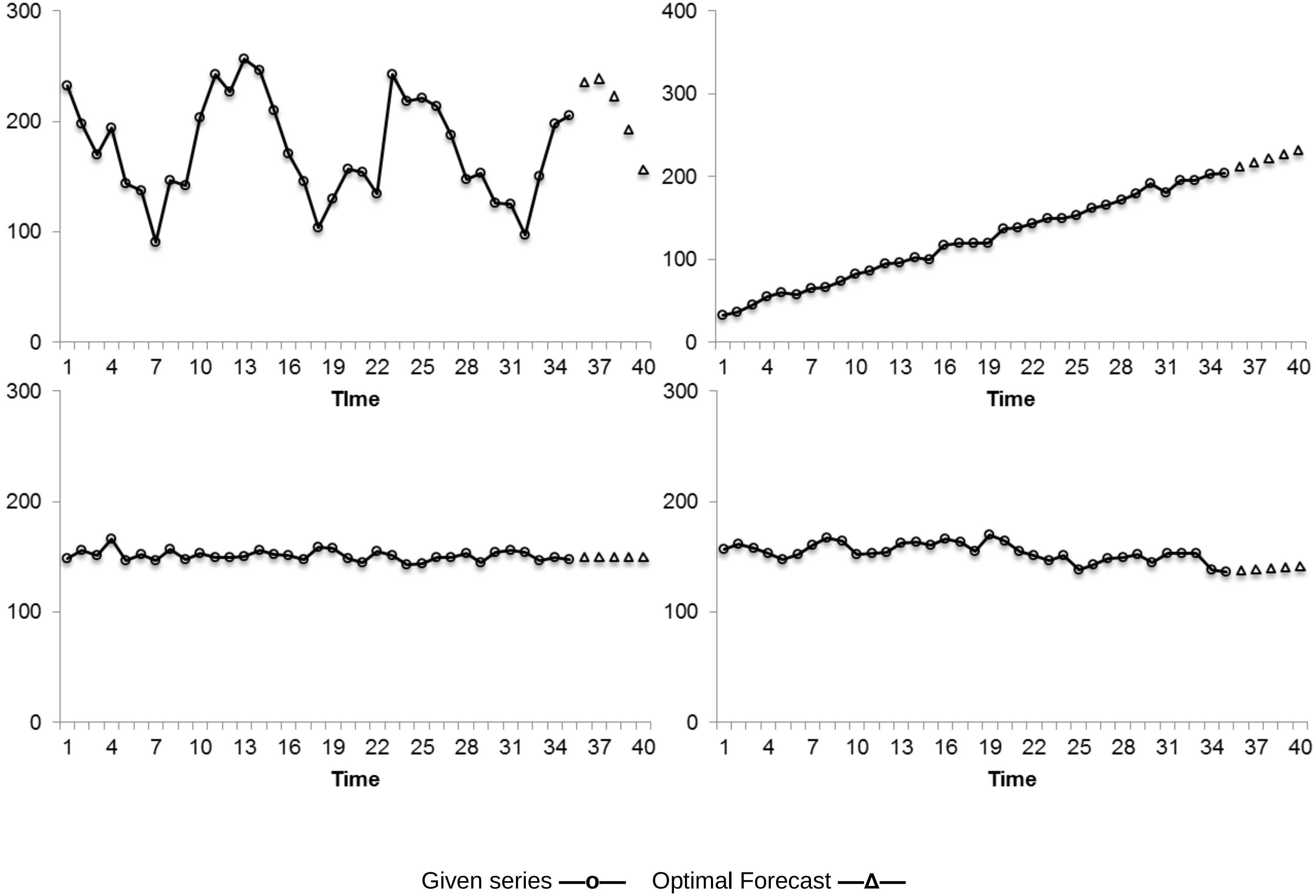
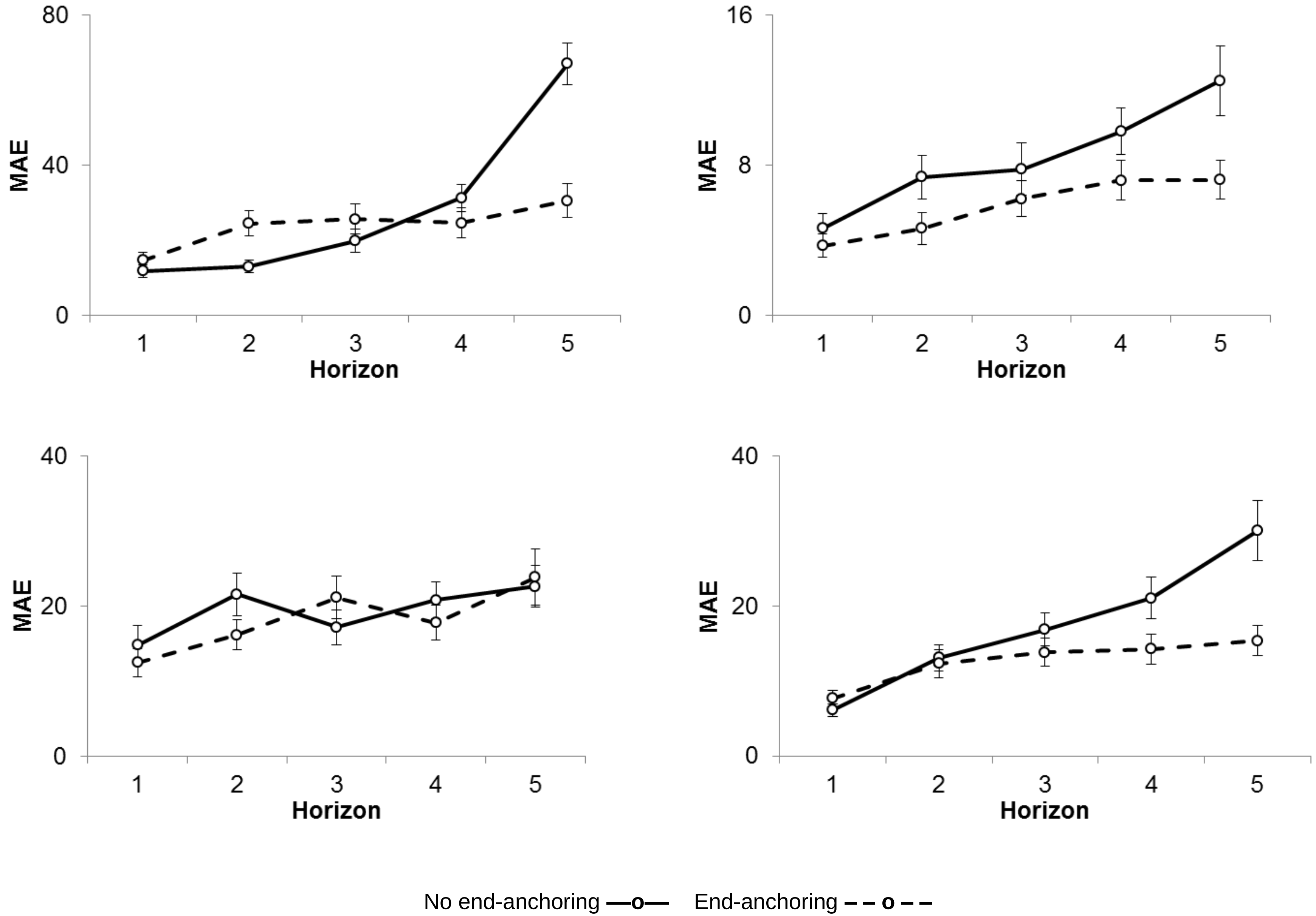

300
200
100
0
1
4
7
10
13
16
19
22
25
28
31
34
37
40
TIme
400
300
200
100
0
Time
Time
Time
Given series —o— Optimal Forecast —Δ—

Figure 2



80
40
0
MAE
1
2
3
4
5
Horizon
16
8
0
MAE
1
2
3
4
5
Horizon
40
20
0
MAE
1
2
3
4
5
Horizon
40
20
0
MAE
1
2
3
4
5
Horizon
No end-anchoring —o—
End-anchoring – – o – –

Figure 3

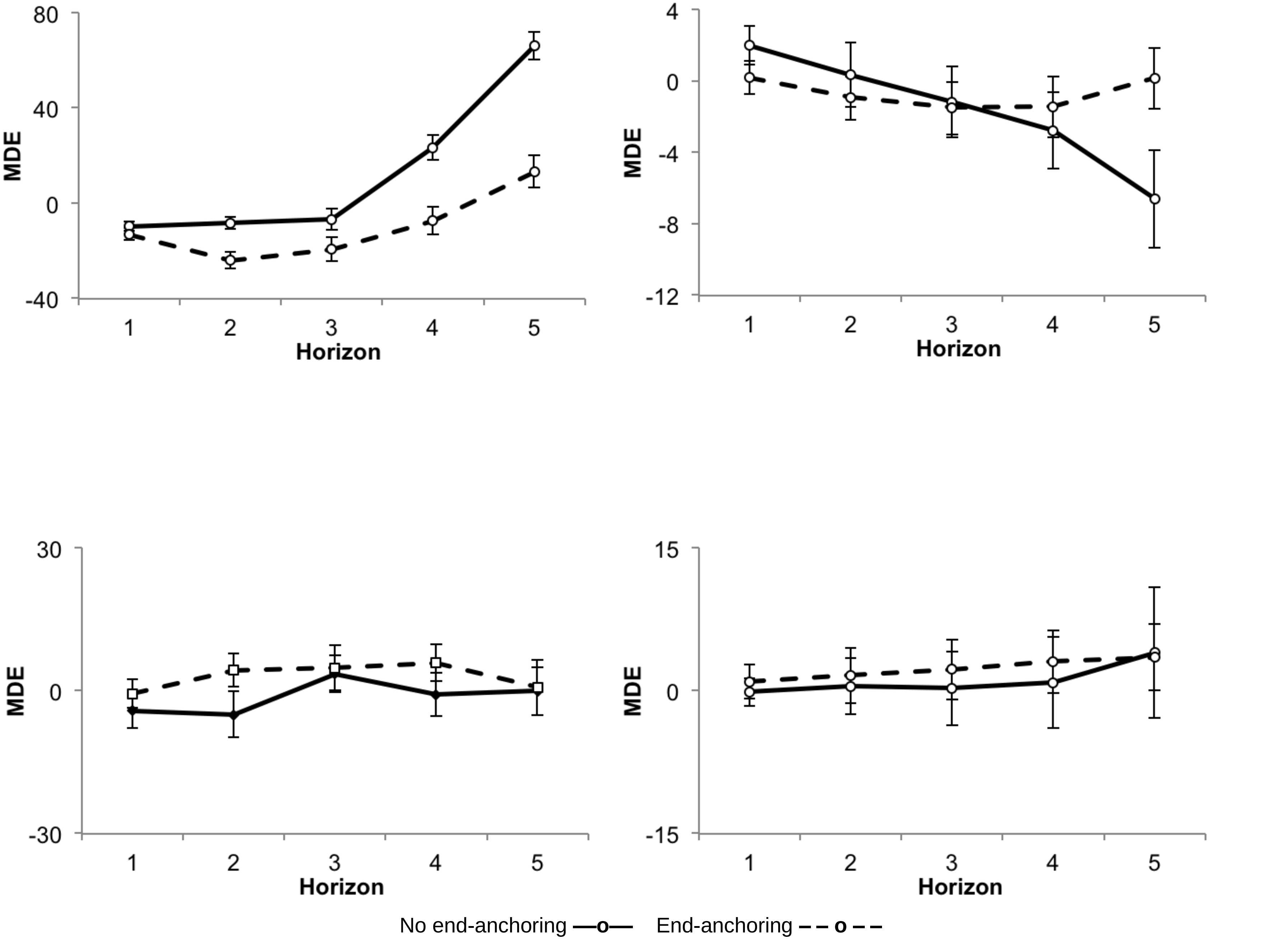

80
40
0
-40
MDE
1
2
3
4
5
Horizon
4
0
-4
-8
-12
MDE
1
2
3
4
5
Horizon
30
0
-30
MDE
1
2
3
4
5
Horizon
15
0
-15
MDE
1
2
3
4
5
Horizon
No end-anchoring —o—
End-anchoring – – o – –

Figure 4

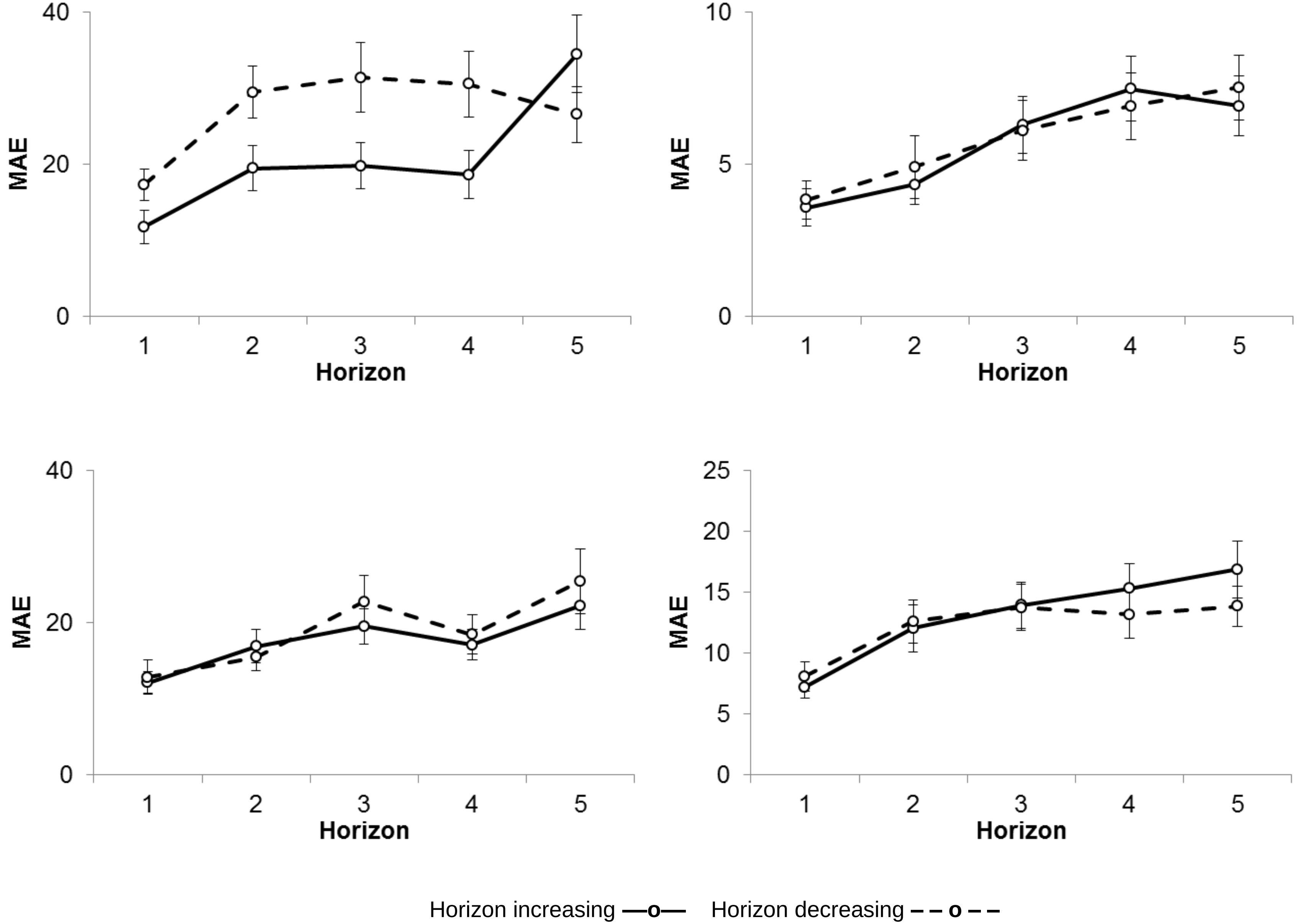

40
20
0
MAE
1 2 3 4 5
Horizon
10
5
0
MAE
1 2 3 4 5
Horizon
40
20
0
MAE
1 2 3 4 5
Horizon
25
20
15
10
5
0
MAE
1 2 3 4 5
Horizon
Horizon increasing —o— Horizon decreasing – – o – –

Figure 5

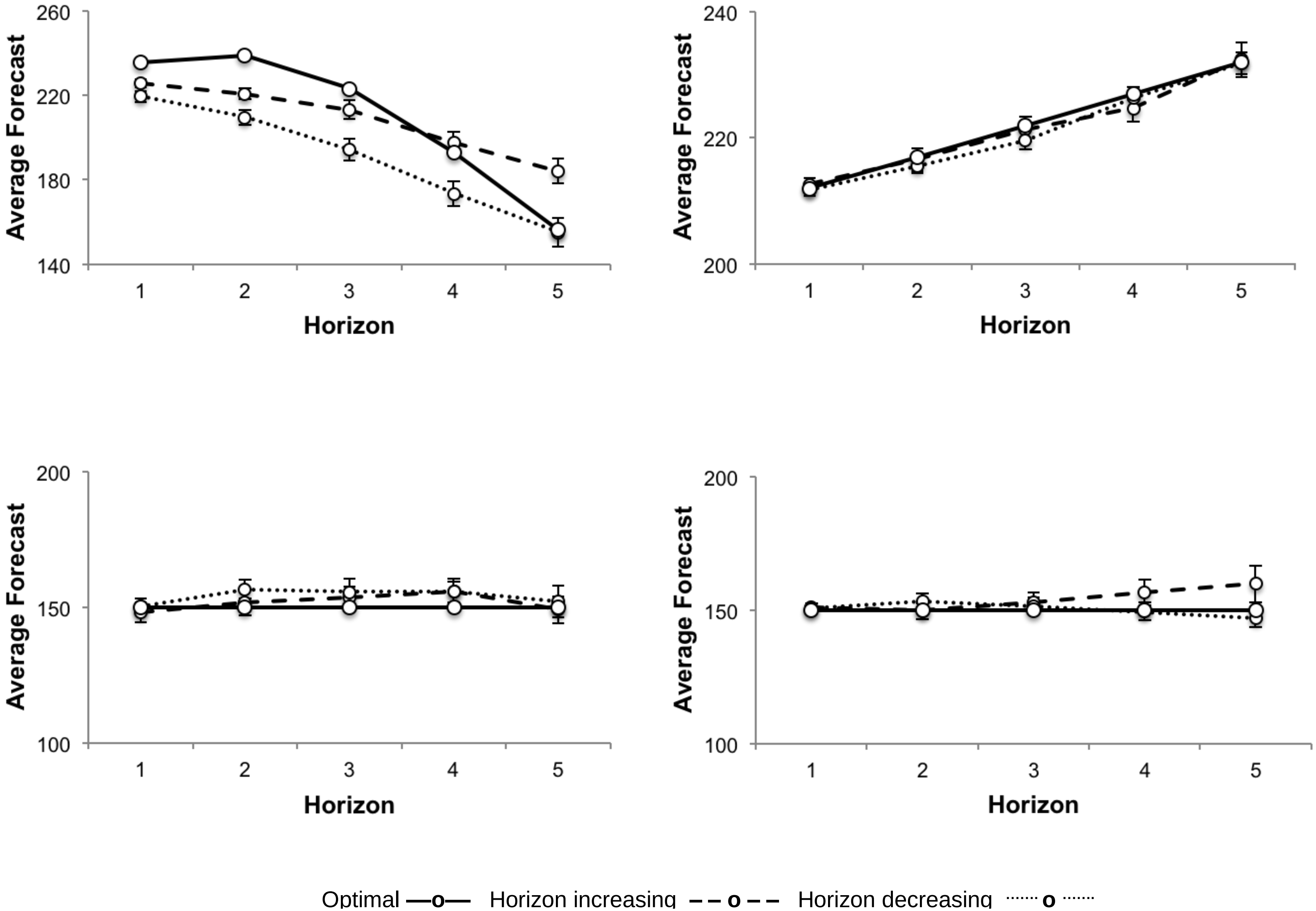
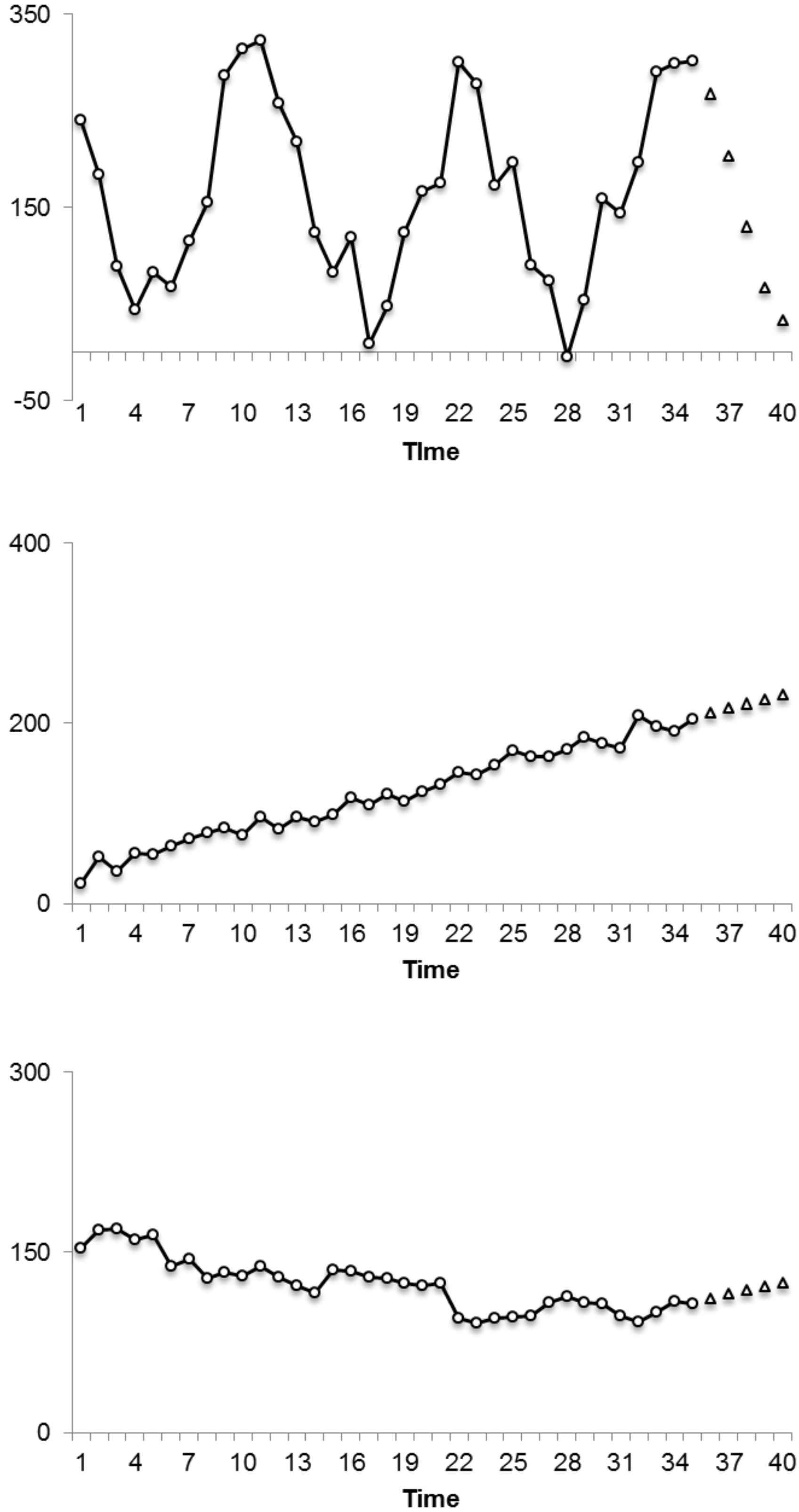
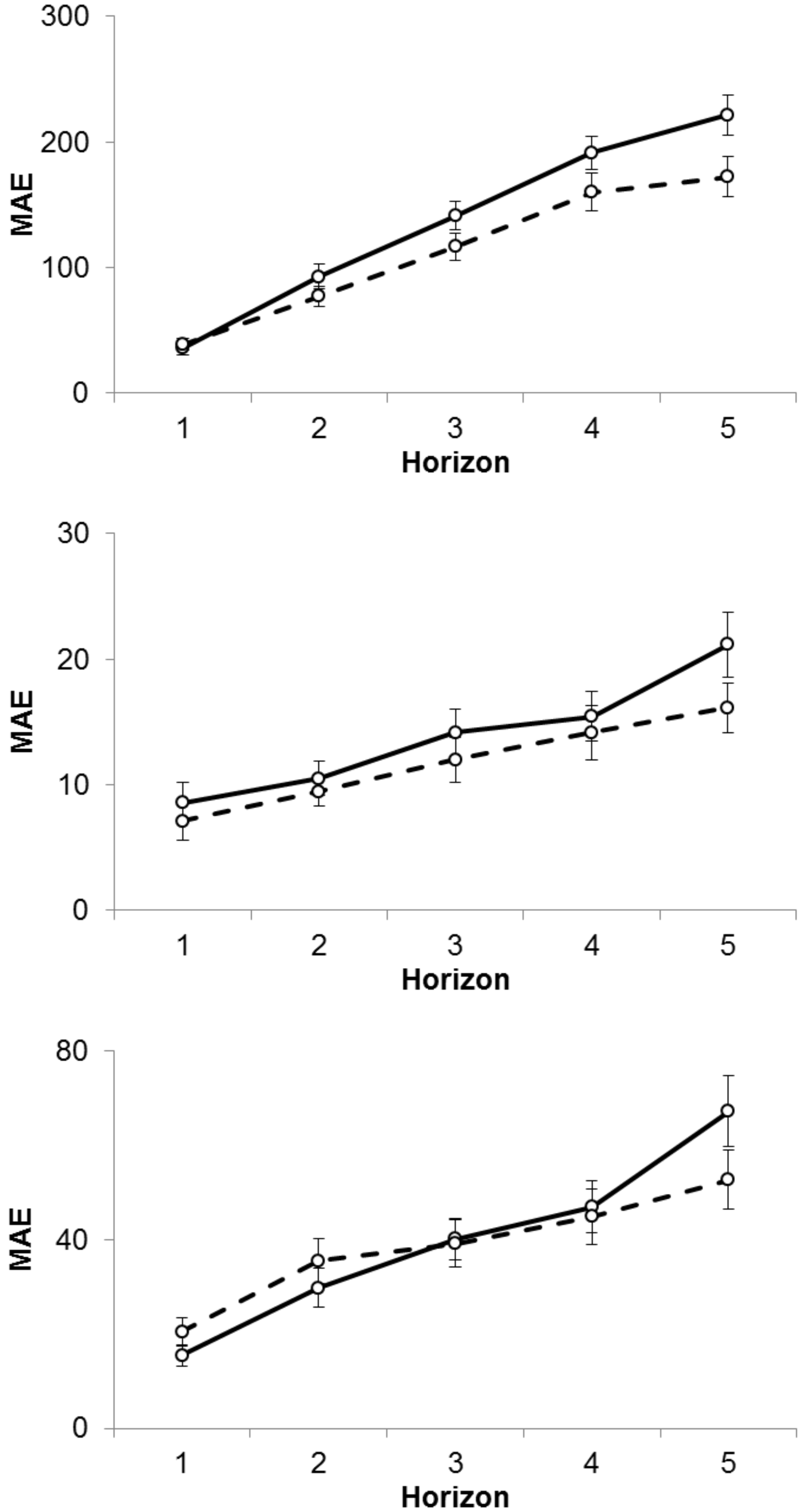

260
220
180
140
Average Forecast
1
2
3
4
5
Horizon
240
220
200
Average Forecast
1
2
3
4
5
Horizon
200
150
100
Average Forecast
1
2
3
4
5
Horizon
200
150
100
Average Forecast
1
2
3
4
5
Horizon
Optimal —o—
Horizon increasing – – o – –
Horizon decreasing ······ o ······

Figure 6



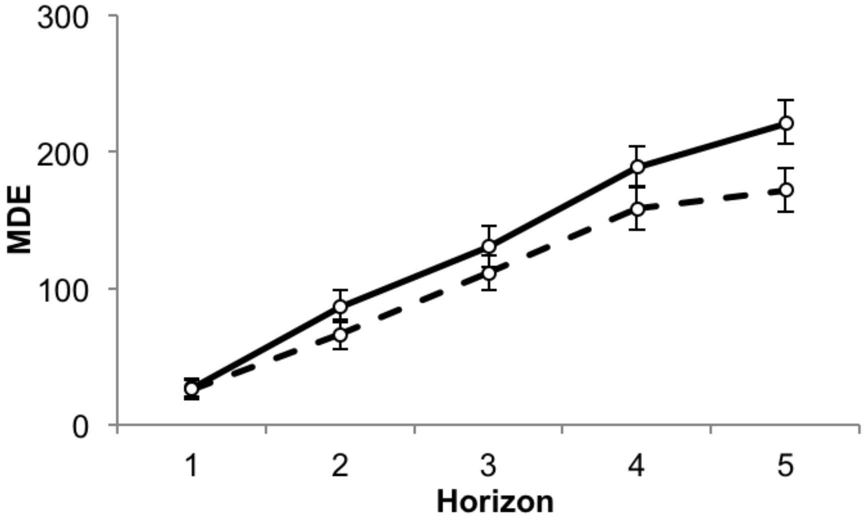
Given series —o— Optimal Forecast —Δ—

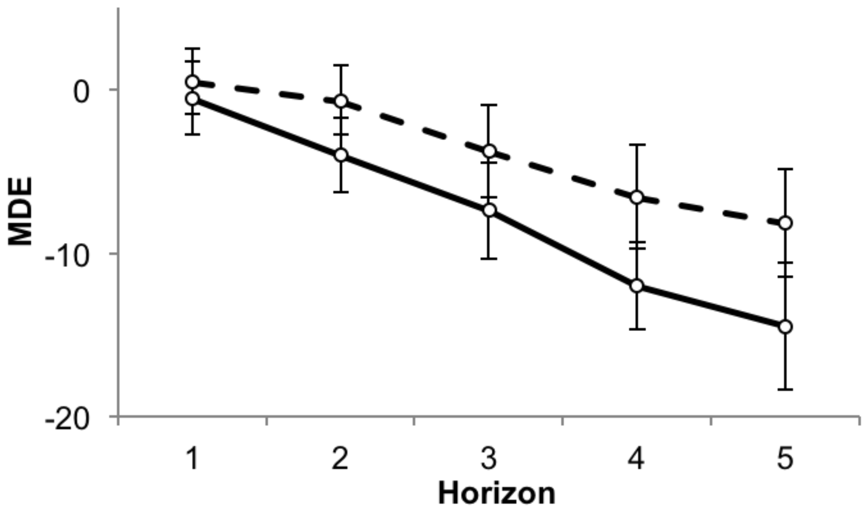
Figure 7

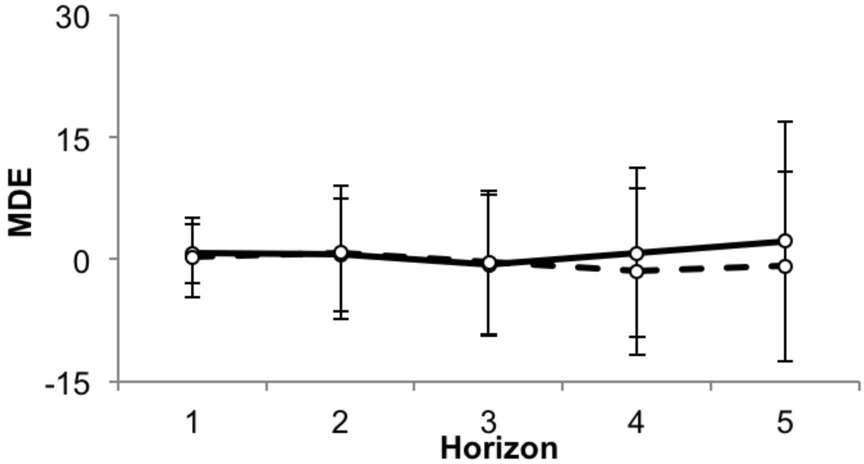


No end-anchoring **—o—** End-anchoring **– – o – –**

Figure 8







No end-anchoring —o— End-anchoring – – o – –

Figure 9

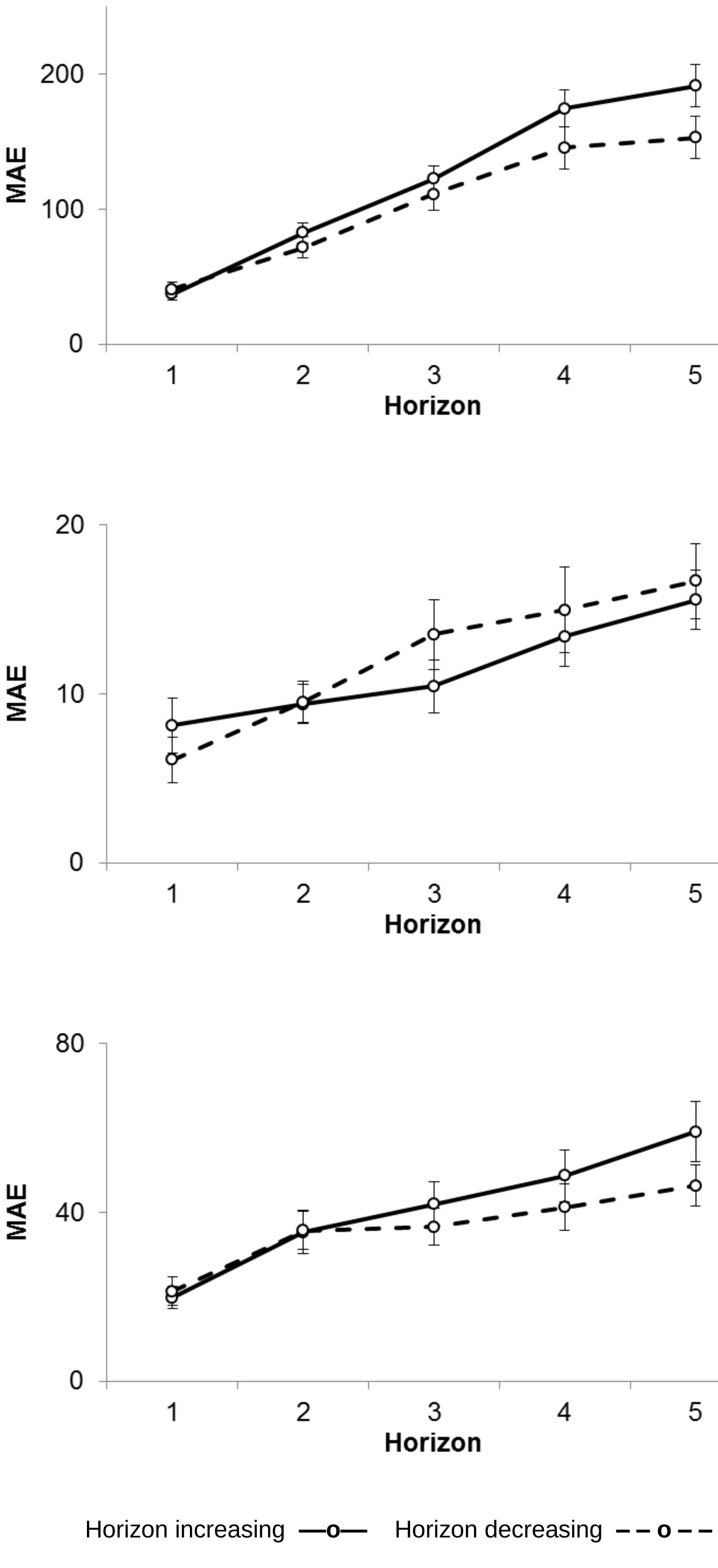
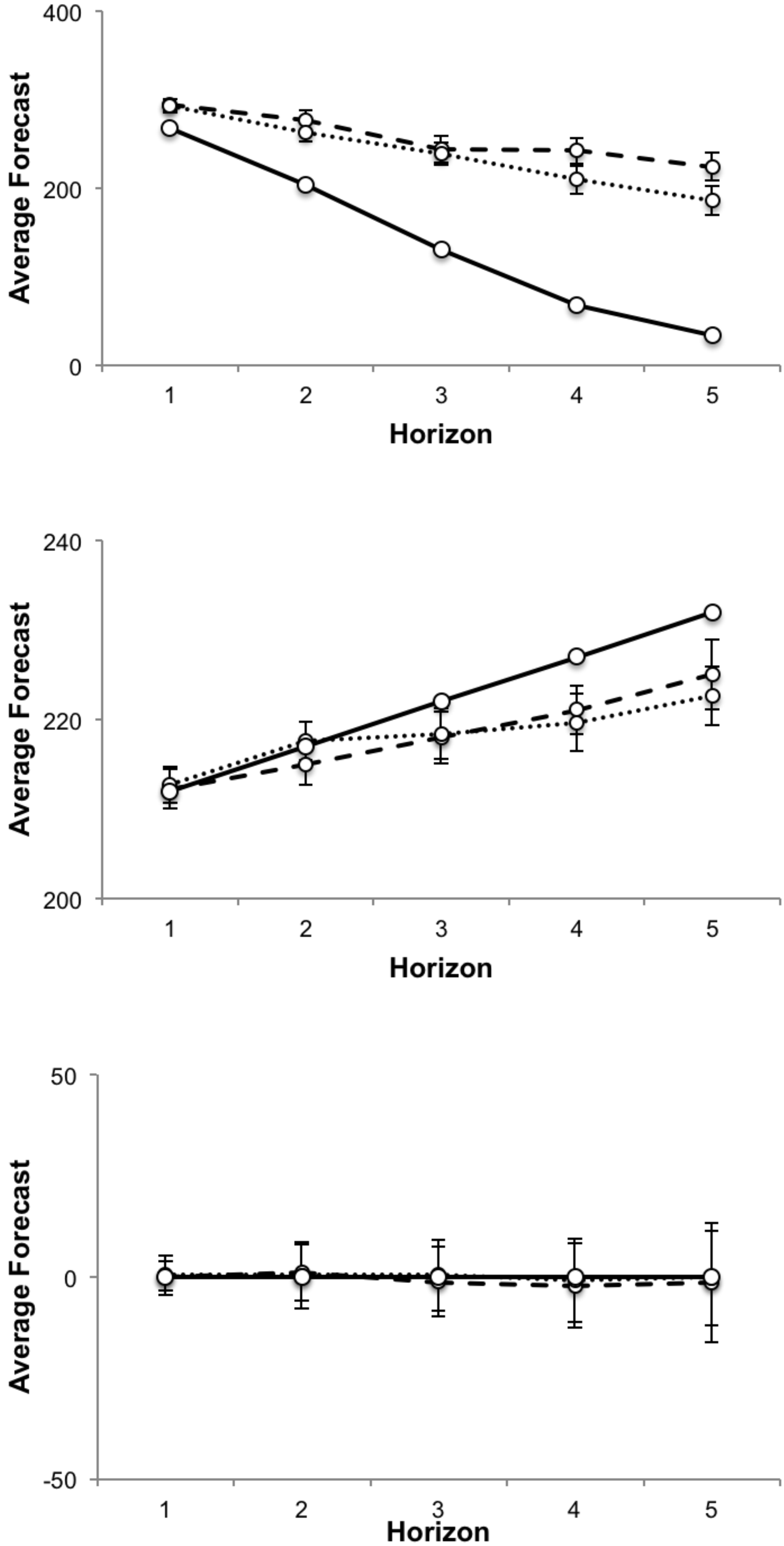

200
100
0
MAE
1
2
3
4
5
Horizon
20
10
0
MAE
1
2
3
4
5
Horizon
80
40
0
MAE
1
2
3
4
5
Horizon
Horizon increasing
Horizon decreasing

Figure 10



Optimal —o— Horizon increasing – – o – – Horizon decreasing ······· o ·······