## Title
Judgmental Time Series Forecasting: A systematic analysis of graph format and trend type



## Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

## Authors
tumen, ceyda
Boduroglu, Aysecan

## Publication Date
2022

Peer reviewed

# Judgmental Time Series Forecasting: A systematic analysis of graph format and trend type

**Ceyda Tumen (ceydatumen@yahoo.com)**
Bogazici University, Department of Psychology, Bebek
Istanbul, Turkey

**Aysecan Boduroglu (aysecan.boduroglu@boun.edu.tr)**
Bogazici University, Department of Psychology, Bebek
Istanbul, Turkey

**Abstract**

In many areas like economics, finance, and health, people make judgmental forecasts looking at previous time series data. In such efforts, either tabular presentations or graphs are utilized, where graphs can be in different formats like bars, lines or points. Different presentations may cause certain biases stemming from bottom-up processing. To delineate such perceptually driven biases in judgmental forecasting, we investigated the effect of graph format (line, bar, point) and trend type (upwards, downwards, flat) on judgmental point forecasts when no domain information was provided. Bringing together perspectives from graph processing, visualization and forecasting literatures, our major goals were to determine which graph formats lead to more accurate forecasts and whether bar graphs lead to mean reversion bias or within-the-bar bias in forecasts. Additionally, we wanted to determine whether asymmetric damping observed in sales forecasts of downward vs. upward trended series were confounded by graph characteristics. We found that forecasts in line and point graphs were less biased than those in bar graphs; forecasts based on bar graphs depicting trended data exhibited mean reversion bias. We also observed a general positivity bias in forecasts for all trend types in line and point graphs. This implied trend following forecasts in upward trends and mean reverting forecasts in downward trends revealing an asymmetricity in the absence of context as well.

**Keywords:** judgmental forecast, graph, trend

## Introduction

Many people engage in tasks that require them to make forecasts based on a given set of past data. Judgmental point forecasts are mostly utilized in the economics domain, as people make future predictions of inflation, stock/fund prices, returns and product sales. Naturally the judgmental forecasting literature has mainly focused on how such financial/economic forecasts vary as a function of domain knowledge, trend type and noise (Bolger & Harvey, 1993; De'Bondt, 1993; Glaser, Iliewa, Weber, 2019; Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997); nevertheless, effects of format characteristics have not been systematically investigated. The graph visualization literature found evidence of differences in bottom-up processing of line, bar and/or point graphs leading to differences in judgments (Correll and Heer, 2017; Godau et al., 2016; Kang et al., 2021; Newman & Scholl, 2012; Schah & Freedman, 2011; Strobel et al., 2016; Xiong et al., 2019; Yuan et al., 2019; Zachs & Tversky, 1999). Different graph formats can possibly lead to biased judgmental forecasts as well (Theocharis, Smith and Harvey, 2019). Our goal was to examine certain biases that typically impact forecasts. Forecasts may be biased either towards the x-axis (within-the-bar bias; Goddau et al., 2016; Kang et al., 2021), the mean of the series leading to mean-reversion (also known as trend-dampening, Bolger & Harvey, 1993; De'Bondt, 1993; Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997) or the last point(s) in the series (recency bias; Glaser et al., 2019; Theocharis et al., 2019). Also, typically forecasts of downward trending series are more mean reverting than forecasts of upward trending series leading to asymmetric damping (O'Connor et al, 1997). We systematically investigated the impact of different graph formats (line, bar and point) and trends (upward, downward and flat) on judgmental forecasts. In the following section, we will first provide a selective review of biases caused by different graph formats, followed by domain and trend effects studied in the judgmental forecasting literature. Then we will provide an overview of the experiment and outline our hypotheses.

**Graph Formats & Judgmental Biases.** Graphs can make statistical information easy to understand, paving the way for easier detection of trends and other patterns in data (Pinker, 1990). People can interpolate and extrapolate functions using the data conveyed via graphs (Ciccione & Dehaene 2021, Schulz et al., 2017) Nevertheless, graphs can mislead people to detect patterns even when none exists (Lawrence et al., 2006). Informationally equivalent graphs may not always be computationally equivalent because visual processing of graphs is impacted by various cognitive heuristics and biases (Shah & Hoeffner, 2002; Zacks & Tversky, 1999). One major bottom-up factor that may cause systematic judgment biases is the graph format (bar, line vs. point graphs). Different graph formats can bias time series forecasting as well (Theocharis et al., 2019).

One typically used graph format when presenting trend data are line graphs. While lines are more suitable in conveying continuous data, bars are preferred for discrete data (Zacks & Tversky, 1999). Accordingly, trends are more



In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. 

accurately extracted from the line graphs. When participants forecast based on data presented in line graphs, responses revealed an effect tagged as recency bias, i.e. forecasts were closer to the last data point compared to when data was presented as point graphs (Theocharis et al., 2019).

When processing bar graphs, a salient bias known as the within-the-bar bias emerges. This leads viewers to mistakenly think the height of the bar represents the likelihood of the values in the distribution rather than just their average (Newman & Scholl, 2012). When there were multiple data points presented via bar graphs, the average estimations were biased more towards the x-axis as a result of the within the bar bias (Goddau et al., 2016; Kang et al., 2021; Yuan et al., 2019, but also see Xiong et al., 2019). Within the bar bias also led to trend line estimates with lower intercepts in bar graphs vs. line and point graphs, implicitly indicating lower perceived means (Correll & Heer, 2017). In an unpublished study, Harvey and Reimers (2012) found that using bars vs. lines or points led to within-the-bar bias in judgmental forecasts, i.e. forecasts were closer to the x-axis both in upward and downward trended graphs. All in all, graph visualization literature associated line graphs with the recency bias and bar graphs with the within the bar bias.

**Effects of Domain Knowledge and Trends.** Researchers detected reversals in sales and price forecasts based on trended series presented in line graphs. (Bolger & Harvey, 1993; De'Bondt, 1993; Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997), a bias called trend-damping. We note that trend damping resembles mean reversion. For instance, Lawrence and Makridakis (1997) associated damping with the general anticipation of reversals in economic time series. Harvey and Reimers (2013) mentioned the adaptation effect, which states that the degree of damping would increase in line with the past knowledge of the world (which usually demonstrates long term-cycles and trends are a part of these cycles) and as the magnitude of the trend slope increases (i.e. damping is not detected in very shallow slopes). In a similar vein, De'Bondt (1993) also showed that financial experts were more likely to predict reversals in trended stock price series, while non-experts' forecasts were more trend following.

Another important finding in these studies was that the reversals were more significant in downward trended series in comparison to upward trended series (Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997). This phenomenon was called the asymmetric damping. There may be several factors contributing to this pattern. First, people are more familiar with rising series than declining ones in real life; this results in higher variance and larger confidence intervals in forecasts of downward trended series (Harvey & Bolger, 1996; Lawrence & Makridakis, 1989, O'Connor et al, 1997). Secondly, the data utilized in these studies were such that higher values were better than lower ones as in sales or stock prices. As a result, an asymmetric damping was observed either due to an optimism bias and/or the perceived likelihood of actions expected to be taken when the data was going down. (Harvey & Reimers, 2013; Lawrence & Makridakis, 1989; O'Connor et al., 1997).

Another line of research focusing on the relationship between forecasting and domain are price-return studies (Glaser et al., 2007; Glaser et al., 2019). In Glaser et al. (2019) participants estimated future values using either prices (via line graphs) or returns of the same investment instruments (via bar graphs). Glaser et al. (2019) found that forecasts were more trend following in the price condition and were more mean reverting in the return condition. Additionally, they found that judgmental forecasts were based on the whole data in the return graphs (i.e. all four quarters of the past data) implying mean reversion and on the more recent data in the price graphs (i.e. only the last quarter's data) implying a recency bias. They attributed these differences to the top-down effects of framing the graphs as prices or returns. However, in this study context was confounded with graph type. Thus, we believe that these results could have been at least partly due to prices being depicted in line and returns being depicted in bar graphs.

In short, during judgmental forecasting tasks, line graphs may lead to a recency bias (Glaser et al., 2019; Theocharis et al., 2019) and bar graphs may lead to a within-the-bar bias (Harvey & Reimers, 2012) or mean reversion (Glaser et al., 2019). Mean reversion is different from within-the-bar bias, since these biases reflect different mechanisms: the first takes into account the mean of the series and the second takes into account the area of the bars pulling the forecasts towards the x-axis as argued by Kang et al, 2021. The two also imply different results specifically in downward trended graphs, i.e. higher estimates (revealing a dampening in the trend) in the case of mean reversion and lower estimates (revealing trend continuation) in the case of within-the-bar bias.

**Goals of study and hypotheses.** In this study, we asked people to make one-period (close) and three-period (far) ahead forecasts, as well as mean estimations based on data presented in different formats. Our first goal was to determine which graph format leads to the highest accuracy, i.e. lowest absolute distance between the judgmental forecasts and the model forecasts. We hypothesized that point graphs would lead to more accurate forecasts than line and bar graphs, as no particular bias unique to point graphs have yet been identified. Our second goal was to determine whether bar graphs lead to mean reversion or within-the-bar bias. We hypothesized that in bar graphs, forecasts would be lower in upward trended series and higher in downward trended series implying mean reversion. Accordingly, the absolute difference from the mean would be lower in both upward and downward trended bar graphs in comparison to line and point graphs. Our third goal was to see whether there would be asymmetric damping in downward vs. upward trended graphs, when no context was specified. We expected forecasts in downward trended graphs to be more mean reverting than upward trended graphs, implying asymmetric damping. We also wanted to explore the mean estimates in terms of whether they are being affected from mean reversion or within-the-bar bias. We used the absolute difference between the mean estimates and the actual mean as our dependent variable. The findings here would shed light on whether forecasts in bar graphs were being affected by the differences in how existing data was mentally summarized. Our final goal was to study whether the line graphs would lead to a recency bias in comparison to bar and point graphs. We hypothesized that the difference between the forecast and the last data point would be smaller in line graphs in all three trend conditions. This would imply recency bias in line graphs not only in flat series (as in Theocharis et al., 2019) but also in trended series (Glaser et al., 2019.

# Method

## Participants

We planned to recruit 40 participants for each of the three graph format groups assuming medium sized effect and 80% power. A total of 199 people entered our Qualtrics link. We excluded those participants who did not complete the experiment (27), entered the experiment via mobiles despite our instructions (15), completed the experiment in less than 4 minutes and over an hour (6), and/or failed the attention check (12). This reduced the sample to 139 participants. Then we chose the first 40 participants for each graph group as planned, resulting in a total of 120 participants (79 females). Data from one participant (from the line group) was excluded from the study and replaced with a new participant because his estimates were 3.3 SD off the mean of his group. Of the 120 participants, 84 were undergraduate students (*Mean age*: 20.71, *SD*: 1.92, *range*: 19-30) and 36 were college graduates (*Mean age*: 41.75, SD: 8.9, *range*: 24-57). The mean age of the overall participants was 27 (*SD*=10.9, *Median*: 21, *range*: 19-57). Undergraduate participants received .5 course credit in exchange for their participation; the rest were reached via snowballing. Study was pre-registered. (AsPredicted #66001)

## Materials

**Judgmental Forecasting Task.** In the judgmental forecasting task, participants were randomly assigned to one of the three graph formats (line, bar, point). Then, after reading the instructions, each participant was shown three graphs presenting a flat, upward and downward trend in random order (see Figure 1). Each graph was presented for two times, first for making one-period ahead and second for three-period forecast. Participants were asked to make their forecasts on a slider scale marked 0-12, where they could choose a decimal number like 6.5. To ensure that content-based expectancies would have minimal impact on the forecasts, we did not provide any title or y-axis label. Before moving to the second part of the experiment, we added an attention check, which asked the participants to move the scale to a point between 70%-80% on a 0-100% scale. In the second part, we asked participants to provide the mean value of the data presented with the exact same graphs again in a randomized order. Each graph was shown for seven seconds at most or until the participant pressed a key to continue, to prevent them from making exact calculations by giving unlimited time. After the graphs disappeared, participants provided their mean estimation on a continuous 0-12 scale, where they could again choose a decimal number.

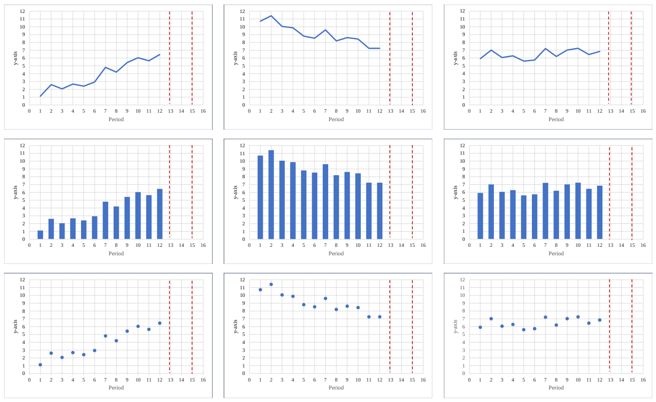

Figure 1: Graphs used as stimuli in the experiment.

The data series used to construct the graphs were made up of 12 points representing the period (as in Glaser et al., 2019). The data was created with a generative model (a+bx+e), in which the slope (b) was 0.4 for upward trended graphs, 0 for flat graphs and -.4 for downward trended graphs and the error term (e) was randomly withdrawn from a normalized distribution with mean 0 and standard deviation .5. Our generation model was similar to Correll et al. (2017) and O'Connor et al. (1997). We manipulated the last data point in the series by varying the error term using either a positive (0.35) or a negative number (-0.35), aiming to control the effect of last change on the judgmental forecasts. Half of the participants in each graph format saw a positive and the remaining half saw a negative last error term. The intercept (a) was determined in a way to make sure that the generative model forecast for the next period was the same for all three trend types. Finally, the graphs were presented with horizontal and vertical grid lines to make data reading easier as in Lawrence and Makridakis (1989).

## Procedure

The experiment was carried online via Qualtrics (https://www.qualtrics.com). Participants provided informed consent and then were randomly assigned to one of the three graph format groups in the judgmental forecasting task. Finally, they completed the Demographic Form and were thanked before they left the site. The study was approved by the Institutional Review Board at Bogazici University.

# Results

We conducted 3x3 mixed factorial design ANOVAs with graph format (line, bar, point) as between participants variable and trend type (upward, downward, flat) as within participants variable, for all dependent variables (the absolute difference between the judgmental forecast and the model forecast, the judgmental forecast and the mean of the series, the judgmental forecast and the last data point, the mean estimate and the mean of the series). As part of the data cleaning, we replaced six data points with sample means as their z-score were below or above 3.3 SD. Greenhouse–Geisser correction was implemented for non-spherical data along with a Bonferroni adjustment for inflated Type 1 error with $\alpha$ set as 0.05.

**Absolute error: Distance from the model forecast.** To determine whether graph format impacts accuracy, we calculated the absolute difference between the judgmental forecast and the model forecast for both horizons. The results for the one-period ahead forecast indicated that the main effect of trend ($F(1.63, 190.8) = 14.69$, $p < .001$, $\eta2 = .11$) was significant. However, there was no main graph effect ($F(2, 117) = .86$, $p = .42$, $\eta2 = .01$) or interaction ($F(3.26, 190.81) = 1.98$, $p = .11$, $\eta2 = .03$). Post-hoc analysis showed that absolute difference was significantly higher in downward trend in comparison to upward trend ($p < .001$) and flat trend ($p < .001$).

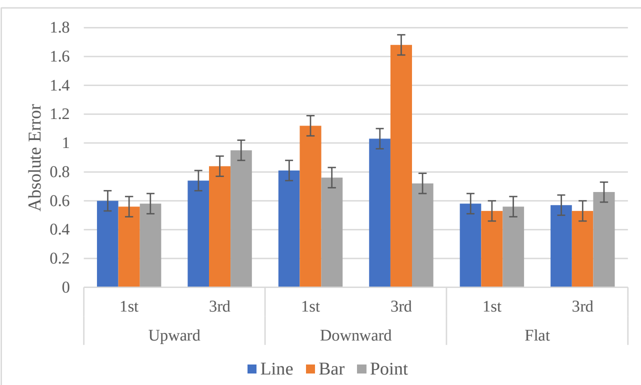


Figure 2: Absolute error for one- and three-period ahead forecasts. Error bars represent standard errors

The analyses for the three-period ahead forecasts implied that there were main effects of trend type ($F(1.60, 186.93) = 16.5$, $p < .001$, $\eta2 = .12$) and graph format ($F(2, 117) = 3.65$, $p < .05$, $\eta2 = .06$). The bar graph was marginally different from both line graph ($p = .07$) and point graph ($p = .06$), as the absolute difference was higher in bar graphs in comparison to line and point graphs. Separately, all trend types were significantly different from each other, where lowest distance from the model forecast was observed in the flat trend, followed by the upward trend ($p <$ =.001) and then the downward trend ($p < .001$). Additionally, there was a significant interaction between graph format and trend type ($F(3.20, 186.93) = 7.13$, $p < .001$, $\eta2 = .11$). The interaction was due to the fact that forecasts in bar graphs were significantly less accurate than point ($p < .001$) and line ($p < .05$) graphs in the downward trended series, but not in upward or flat series. All in all, in the bar graph condition forecasts were less accurate, specifically for the downward trended series. There was, however, no significant accuracy difference between line and point graphs regardless of trend type.

**Mean reversion or within the bar bias: Distance from the Mean.** We calculated the absolute difference between the forecast and the mean of the series to determine whether there was mean reversion or within-the-bar bias. For the one-period ahead forecasts, both the main effect of graph format ($F(2, 117) = 10.56$, $p < .001$, $\eta2 = .15$) and trend type ($F(1.59, 185.94) = 330.51$, $p < .001$, $\eta2 = .74$) were significant. However, there was no interaction effect ($F(3.18, 185.94) = 1.77$, $p < .001$, $\eta2 = .03$). Post-hoc analysis showed that distance from the mean (descriptive statistics presented in Table 1) in the bar graph condition was significantly smaller compared to line ($p < .001$) and point graphs ($p < .001$) and this was valid for all trend types ($p < .001$) indicating relative mean reversion (and not within-the-bar bias). There was no significant difference between point and line graph conditions. Separately, all trend types were significantly different from each other ($p < .001$), where lowest distance from the mean was seen in the flat trend, followed by the downward trend and then the upward trend. This indicated existence of asymmetric damping between upward and downward trended series.

Table 1: Descriptive statistics for the absolute difference between forecasts and actual mean

| | | 1-period | | 3-period | |
|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD |
| Upward | Line | 3.00 | .70 | 3.94 | .82 |
| | Bar | 2.42 | .64 | 3.04 | .99 |
| | Point | 2.92 | .66 | 3.75 | 1.06 |
| Downward | Line | 2.03 | .83 | 3.11 | 1.22 |
| | Bar | 1.70 | .67 | 2.22 | 1.39 |
| | Point | 2.04 | .89 | 3.09 | .98 |
| Flat | Line | .60 | .45 | 0.59 | .45 |
| | Bar | .54 | .36 | 0.54 | .44 |
| | Point | .60 | .46 | 0.68 | .46 |

The analyses for the three-period ahead forecasts mirrored those of the one-period ahead forecasts. As before, there were main effects of trend type ($F(1.70, 198.94) = 378.36$, $p < .001$, $\eta2 = .76$) and graph format ($F(2, 117) = 13.08$, $p < .001$, $\eta2 = .18$). The bar graph was significantly different from both line graph ($p < .001$) and point graph ($p < .001$), as distance from the mean was smaller in bar graphs in comparison to line and point graphs. This was valid for both upward and downward trended graphs ($p < .001$), indicating relative mean reversion. Also, all trend types were significantly different from each other ($p < .001$), where lowest distance from the mean was registered by the flat trend type, followed by the downward and upward trend types, again implying asymmetric damping. This was observed across all graph formats (($p < .001$). Additionally, there was a small yet significant interaction between graph format and trend type

($F$(3.4, 198.94) = 3.56, $p < .05$, $\eta 2 = .05$). This was due to the fact that bar graphs were more mean reverting than the other graph formats in upward and downward trended series, while there was no significant difference in the flat series.

**Mean Estimate Accuracy: Distance between the Actual Mean and the Mean Estimate.** To explore if differences in forecasts are caused by differences in mean estimates, we calculated the absolute difference between the actual mean and the mean estimate. There was a significant main effect of trend type ($F$(2,234) = 4.85, $p < .01$, $\eta 2 = .04$). Flat graphs were significantly different from both upward trended ($p < .05$) and downward trended ($p < .05$) graphs as the distance between the actual mean and the mean estimate was lower in flat graphs ($M$ = 1.18, $SD$ = 1.04) with respect to upward ($M$ = 1.57, $SD$ = 1.40) and downward trended graphs ($M$ = 1.55, $SD$ = 1.29). As no surprise, the higher accuracy in flat graphs imply a relative ease while estimating means using flat graphs compared to trended graphs. What is more important is that there was no significant difference between mean estimation accuracy of upward and downward graphs. There was neither main effect of graph type ($F$(2,117) = 2.80, $p = .08$, $\eta 2 = .04$) nor interaction effect between graph format and trend type ($F$(4,234) = 1.72, $p$ =.15 , $\eta 2 = .03$). Thus, the mean reversion bias observed in forecasts in the bar graph condition or the asymmetricity between forecasts of upward and downward trends cannot be attributed to how viewers mentally summarized studied trends.

Table 2: Descriptive statistics for the absolute difference between one period ahead forecast and last data point

| | | Mean | SD |
|---|---|---|---|
| Upward | Line | .84 | .49 |
| | Bar | .52 | .33 |
| | Point | .81 | .56 |
| Downward | Line | .66 | .45 |
| | Bar | .84 | .84 |
| | Point | 86 | .51 |
| Flat | Line | .59 | .38 |
| | Bar | .57 | .39 |
| | Point | .66 | .36 |

**Recency Bias: Distance from the Last Data Point.** To determine whether there was a recency bias in forecasts, we calculated the absolute difference between the forecast and the last data point of the series. There were small yet significant effects of trend type ($F$(1.83, 214.36) = 4.56, $p < .05$, $\eta 2 = .04$) and an interaction between trend type and graph format ($F$(3.66, 214.36) = 2.98, $p < .05$, $\eta 2 = .05$). There was no effect of graph format ($F$(2, 117) = 1.70, $p = .19$, $\eta 2 = .03$). Flat graphs were significantly different from downward trended graphs ($p < .05$) and marginally different from the upward trended graphs ($p = .06$) as the distance from the last data point was smaller in flat graphs with respect to the downward trended graphs (descriptive statistics presented in Table 2). The interaction was driven by the upward trended bar graphs. They were significantly different from upward trended lines ($p < .01$) and upward trended points ($p < .05$). In fact, the distance from the last data point was smaller in upward trended bar graphs in comparison to lines and points. These findings implied a recency bias in upward trended bar graphs with respect to upward trended line and point graphs. However, the effects were small. Contrary to our expectations, our data did not reveal a recency bias in the line graph condition.

## Discussion

We studied the effects of graph format and trend type on forecasting controlling for top-down domain effects. The novelty of our study was its context-free setting to delineate the bottom-up effects, since forecasting had been mainly studied in context-rich settings (Harvey & Reimers, 2013; Lawrence & Makridakis, 1989; O'Connor et al., 1997). We found that point and line graphs pave the way for more accurate forecasts with smaller absolute error. Bar graphs lead to more mean reverting forecasts with relatively higher damping in both upward and downward trended graphs in comparison to line and point graphs (as shown in Figure 3). This finding indicated that bar graphs caused mean reversion rather than within- the-bar bias, since the first required a higher but the latter required a lower forecast in the downward trend condition (Kang et al., 2021). This finding was in line with our hypothesis and Glaser et al. (2019)'s finding but different from Harvey and Reimers (2012)'s claim for within-the-bar bias in forecasts of both upward and downward trended bar graphs. Critically, we also found an asymmetry in the extend of mean reversion in the forecasts of downward as opposed to upward trended series (which was valid for all graph formats) even when no particular context was specified. This was in line with our hypothesis and replicated earlier findings (Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997), but extended those to forecasting in a no-context case. These findings were observed for both forecasting horizons, displacing the possibility that they were valid only for the short or longer-term.

As can be seen in Figure 3, there was a general positivity bias in forecasts in almost all conditions: the forecasts were above the model forecasts. One exception was the upward trending bar graphs. This resulted in trend continuing forecasts (which were above model forecasts) in upward trending series in line and point graphs, but mean reverting forecasts (which were again above model forecasts) in all three graph formats. This pattern resulted in the asymmetricity in forecasts of upward vs. downward trends. Interestingly, this finding contradicted with Ciccione and Deahene (2021). They had participants make forecasts based on periodic functions similar to ours (i.e. a+bx+e with x being the period and e the error term), which were presented in scatterplots, where no context was specified. They found that forecasts were always trend-following: Specifically, they were above the model forecasts in upward trended series and below the model forecasts in downward trended series in line with the predictions of Deming regressions rather than OLS regressions. The difference between our study and theirs may have stemmed from the fact that we labelled the x-axis in our graphs as "period" and they did not, making their graphs look like scatterplots depicting a relationship between two variables. Our choice might have led some participants to imagine a positive variable such as sales while forecasting, mimicking what they are more likely to encounter in everyday life. This was presumably why our finding of a more pronounced mean reversion in downward as opposed to upward trends resembled the prior findings of asymmetrical dampening witnessed in upward vs. downward sales forecasts (Lawrence & Makridakis, 1989; O'Connor et al., 1997). Differences in how viewers mentally summarized trends does not seem to contribute to the biases observed. Specifically, we cannot explain relative mean reversion in forecasts using bar graphs and the asymmetric mean reversion between forecasts in upward and downward graphs via differences in mean estimates. Rather, our findings seem to be related with the differences in forecasting processes in bar graphs compared to line and point graphs.

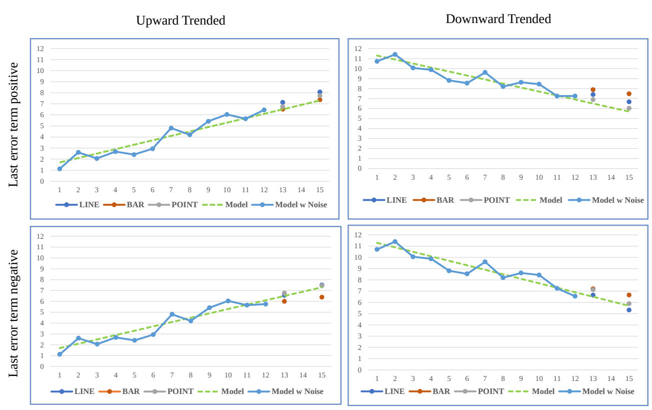


Figure 3: One- and three-period ahead mean judgmental forecasts. Top and bottom panel depicts forecasts based on trends with positive and negative last error term.

Contrary to our expectations, we observed no recency bias in forecasts based on line graphs in comparison to bar and point graphs. This was in contrast to earlier findings by Theochrais et al (2019) and Glaser et al. (2019). Theocharis et al. found recency bias in line graphs in comparison to point graphs using flat and serially independent time series, which resembled our flat series. They linked their finding with an illusory bias of serial dependence caused by line graphs in comparison to point graphs. As our trended data was already made up of serially dependent series, this explanation does not hold for our data. We did find that forecasts in upward trended bar graphs displayed a recency bias in comparison to line and point graphs. However, the effects were small. This may be traced to the participants' different strategies of forecasting in upward trending bar graphs compared to upward trending line and point graphs, which also resulted in mean reverting forecasts in bar graphs vs. trend following forecasts in line and point graphs.

Our findings altogether underline the importance of bottom-up factors such as the different graph formats and trend types in affecting the judgmental forecasts. We argue that the judgmental forecasting literature may have overlooked the impact from bottom-up factors, instead focusing on framework effects, specifically domain knowledge, experience and trend characteristics. One important example is Glaser et al. (2019)'s findings of mean reverting return forecasts vs. trend following price forecasts observed for the same investment instruments. They associated this with the top-down framework effects. Nevertheless, we argue that two bottom-up factors may have played a key role in their results. The first one was that they used bar graphs to convey returns and line graphs to convey prices. We already showed that bar graphs cause relative mean reversion in comparison to line graphs. Additionally, returns were calculated using the prices and this led to visually flatter graphs as per characteristics of returns, whereas price graphs were trended. We also showed that flat graphs lead to more mean reverting forecasts than trended graphs.

One major question that remains to be answered is whether an economic/financial domain leads to trend damping (i.e. mean reversion) in forecasts and a further asymmetry between forecasts of upward vs. downward trended series as suggested by the judgmental forecasting literature (Bolger & Harvey, 1993; De'Bondt, 1993; Glaser et al., 2019; Harvey & Reimers, 2013; Lawrence & Makridakis, 1989, O'Connor et al, 1997). In ongoing work, we compare different domains (no-domain vs. sales), this time controlling for graph format effects using line graphs as in sales forecasting studies mentioned previously. Preliminary findings suggests that the results for the no-domain line graphs replicated findings from our initial study. We believe that this approach that takes into account both bottom-up and top-down factors in judgmental forecasts is likely to give a more comprehensive understanding of biases shaping forecasts.

## References

Bolger, F., & Harvey, N. (1993). Context-sensitive heuristics in statistical reasoning. *The Quarterly Journal of Experimental Psychology Section A, 46(4)*, 779-811.

Ciccione, L., & Dehaene, S. (2021). Can humans perform mental regression on a graph? Accuracy and bias in the perception of scatterplots. *Cognitive Psychology*, 128, 101406.

Correll, M., & Heer, J. (2017, May). Regression by eye: Estimating trends in bivariate visualizations. *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1387-1396).

De Bondt, W. P. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting, 9(3)*, 355-371.

Glaser, M., Langer, T., Reynders, J., & Weber, M. (2007). Framing effects in stock market forecasts: The difference between asking for prices and asking for returns. *Review of Finance, 11(2)*, 325-357.

Glaser, M., Iliewa, Z., & Weber, M. (2019). Thinking prices versus thinking about returns in financial markets. *The Journal of Finance, 74(6)*, 2997-3039.

Godau, C., Vogelgesang, T., & Gaschler, R. (2016). Perception of bar graphs–a biased impression?. *Computers in Human Behavior, 59*, 67-73.

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. (2007). Use of heuristics: Insights from forecasting research. *Thinking & Reasoning, 13(1)*, 5-24.

Harvey, N., & Reimers, S. (2012). Bars, lines, and points: The effect of graph format on judgmental forecasting. *In Proceedings of the 32nd Annual International Symposium on Forecasting.*

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(2)*, 589.

Kang, H., Ji, J., Yun, Y., & Han, K. (2021). Estimating Bar Graph Averages: Overcoming Within-the-Bar Bias. i-*Perception, 12(1)*, 2041669520987254.

Lawrence, M., & Makridakis, S. (1989). Factors affecting judgmental forecasts and confidence intervals. *Organizational Behavior and Human Decision Processes, 43(2)*, 172-187.

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(3)*, 493-518.

Newman, G. E., & Scholl, B. J. (2012). Bar graphs depicting averages are perceptually misinterpreted: The within-the-bar bias. *Psychonomic Bulletin & Review, 19(4)*, 601-607.

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(3)*, 165-176.

Pinker, S. (1990). A theory of graph comprehension. *Artificial Intelligence and the Future of Testing*, 73-126.

Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive psychology, 99*, 44-79.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14(1)*, 47-69.

Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in Cognitive Science, 3(3)*, 560-578.

Strobel, B., Sass, S., Lindner, M. A., & Köller, O. (2016). Do graph readers prefer the graph type most suited to a given task? Insights from eye tracking. *Journal of Eye Movement Research, 9(4)*, 1-15.

Theocharis, Z., Smith, L. A., & Harvey, N. (2019). The influence of graphical format on judgmental forecasting accuracy: Lines versus points. *Futures & Foresight Science, 1(1)*, e7.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5(4)*, 297-323.

Xiong, C., Ceja, C. R., Ludwig, C. J., & Franconeri, S. (2019). Biased Average Position Estimates in Line and Bar Graphs: Underestimation, Overestimation, and Perceptual Pull. *IEEE Transactions on Visualization and Computer Graphics, 26(1)*, 301-310.

Yuan, L., Haroz, S., & Franconeri, S. (2019). Perceptual proxies for extracting averages in data visualizations. *Psychonomic Bulletin & Review, 26(2)*, 669-676.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition, 27(6)*, 1073-1079.