

## **The effect of price volatility on judgmental forecasts: The correlated response model**

**Daphne Sobolev**

### **Abstract**

Traders often employ judgmental methods when making financial forecasts. To characterize judgmental forecasts from graphically presented time series, I propose the correlated response model, according to which properties of judgmental forecasts are correlated with properties of the forecasted series. To test the model, participants were presented with graphs depicting synthetic price series. In Experiment 1, participants were asked to make point forecasts for different time horizons. Participants could control the graphs' time scale. In Experiment 2, participants made multi-period forecasts. They could apply moving average filters on the graphs. Dispersion of point forecasts between participants (the standard deviation of participants' point forecasts) and variability of individual participant's multi-period forecasts (local steepness and oscillation) were extracted. Both forecast measures were found to be significantly correlated with variability measures of the original, scaled, and smoothed data graphs. Thus, the results supported the correlated response model and provided insights into forecasting processes.

**Key words: trading; financial decisions; forecasts; dispersion; horizon; fractal; Hurst exponent; scaling; moving average.**

**Print:** black-and-white (the paper does not contain any color figures. Black-and-white figures are submitted in a separate file).

## 1. Introduction

A high percentage of market participants base their trades on methods which involve extrapolation and pattern recognition of graphically presented financial time series (Batchelor, 2013; Batchelor and Kwan, 2007; Cheung and Chinn, 2001; Taylor and Allen, 1992). Furthermore, it was found that technical analysis techniques are incorporated in decision making processes of the majority of FX dealers (Gehrig and Menkhoff, 2006) and fund managers (Menkhoff, 2010). Nevertheless, the dependence of forecasts from graphically displayed price series on the properties of the data series has not been studied within Finance and has been understudied within Judgmental Forecasting. In particular, the way data graphs' properties affect forecast dispersion (the extent to which forecasters disagree about their forecasts) and forecast variability (the local steepness and oscillation of individual forecaster's multi-period forecasts) has not been explored.

The aim of this paper is to understand the way properties of graphically presented time series affect forecast variability and dispersion. I suggest that the variability of the given time series is correlated with the forecast dispersion of point forecasts and the variability of multi-period forecasts.

Moreover, this effect is robust across different time series, forecast horizons or multi-period forecast densities, and when the forecasters are given the option to scale or smooth the graphs. To provide a theoretical justification for this relation, I propose the correlated response model, described in Section 1.1. The experimental hypotheses are described in Section 1.2.

### **1.1. The correlated response model: Background and definition**

During the past twenty years, a large body of research about the way people make forecasts from graphically presented time series has accumulated. Harvey (1995) showed that when people make multi-period forecasts from graphically presented time series, they tend to imitate the noise component of the time series. The noise level of the forecasts was correlated with the noise level of

the data. Bolger and Harvey (1993) hypothesized that people imitated noise in order to make their forecasts representative of the data series. Furthermore, Harvey, Ewart and West (1997) showed that participants had a strong tendency to imitate the noise component of the data. In one of their experiments, the following instructions were given (page 126): “Put six crosses on the graph to show us your forecasts. Obviously you cannot be certain where these future points will be but try to ensure that your forecasts show the *most likely* positions for them. For example, if you feel that a particular point could lie within a range of values, put your cross in the centre of that range if you feel that this is the most likely position for the true point within the range. Your aim is to maximize the probability that your forecasts will be correct.” Nevertheless, participants in their experiment imitated the noise of the data series.

Lawrence and Makridakis (1989) showed that, though people tend to damp trends, judgmental forecasts correspond to the slope of the given data. Similar results were obtained in other studies (e.g. Bolger and Harvey, 1993). A comprehensive survey about the influence of data characteristics on forecasts was written by Lawrence, Goodwin, O’Connor and Önkal (2006).

Forecast dispersion has not been the topic of studies in Judgmental Forecasting. However, Reimers and Harvey (2011) studied the effect of random noise on judgmental forecasts and mentioned that their experiment “shows that the participants were more variable in their responses when the noise was higher” (see the Result Section of Experiment 1, page 1202 in Reimers and Harvey, 2011). The same result was replicated in their second experiment. Similar relationship between data variability and forecast dispersion was observed for the case of inflation forecasts by Cukierman and Wachtel (1979; 1982). That may be because noisy data is characterized by a high variability and uncertainty. Therefore, it is likely to enable the expression of individual differences more than data with low variability levels. Indeed, in contexts other than forecasting, Caspi and Moffitt (1993) suggested that “individual differences are most likely to be accentuated by unpredictability, when there is a press to behave but no information about how to behave adaptively. Such transition situations are revealing

because during these periods [...] individuals must summon their resources.” Drawing on Caspi and Moffitt (1993), Yang (2012) contended that “individual differences are accentuated when individuals face ambiguous and uncertain events with insufficient information to allow adaptive behaviour”. As highly variable data emphasizes individual differences more than data characterized by low variability, it is, in particular, likely to highlight individual differences in forecasting. However, the latter is expected to result in large group forecast dispersion. Thus, differences in the expression of individual differences in forecasting may explain the relationship between the variability of the given data and forecast dispersion. This explanation is in line with Cukierman and Wachtel (1979), who suggested that differences in the interpretation of volatile data affect forecast dispersion.

Each of the experimental forecasting papers mentioned above examined a highly specialized aspect of forecasting and, thus, contributed to the understanding of forecasting biases and errors.

However, their results have not been united into a single model. Beyond identification of biases, these papers show that, to different levels of accuracy, judgmental forecasts from graphically presented series preserve properties of the given data series. A possible explanation of this preservation may be that imitation is one of the most powerful human learning processes (Bandura and Barab, 1971). In contexts other than forecasting, it was shown that people have an innate tendency to imitate stimuli (Heyes, 2011).

Uniting the results of the experimental papers mentioned above, and generalizing them further, I suggest the correlated response model, formulated below.

**The correlated response model.** In judgmental forecasting tasks, which involve forecasts from graphically presented time series, people’s responses are correlated with properties of the given series. In particular:

1. The trend of the data series and the trend of the forecast are positively correlated (Lawrence and Makridakis, 1989; Bolger and Harvey, 1993).

2. The variability of multi-period forecasts and the variability of the data series are positively correlated (Harvey, 1995; Bolger and Harvey, 1993; Harvey, Ewart, and West, 1997).
3. The forecast dispersion of single point forecasts is correlated with the variability of the data (Reimers and Harvey, 2011).

In this study, the main measure of forecast dispersion of single point forecasts is the standard deviation of the point forecasts made by independent forecasters. (Two other measures of forecast dispersion are examined, too, as described in Section 2.2.1). Two main data variability measures are used: local steepness and oscillation. The local steepness of a graph is defined as the average of the absolute value of the gradients of the graph. The graphs' oscillation is defined as the difference between the maximum and minimum values of the graph over a given interval (Trench, 2002).

As this paper aimed at understanding the way properties of graphically presented time series affect forecast variability measures, I concentrated on Sections 2 and 3 of the correlated response model.

## **1.2 Hypotheses**

The tasks in the experimental studies described in Section 1.1 were not designed to simulate financial situations. For instance, the contexts of the experimental tasks of Reimers and Harvey (2011) were sales forecasts (Experiment 1) and profit forecast (Experiments 2 and 3). In addition, the number of data points presented to the participants in each time series was 50, whereas financial data is abundant and complex. Furthermore, both moving average filters and scaling are commonly offered as options in financial data analysis programs. For instance, Yahoo! Finance (<http://finance.yahoo.com/>) enables users to choose the time scale on which data would be presented and to present different technical indicators, including the moving window average. Moving window averaging filters are applied to smooth graphs. Investors and traders use these options frequently (Glezakos and Mylonas, 2003). However, the experimental settings used by

Reimers and Harvey (2011) did not allow participants to change the scale of the presented data or to smooth it. Finally, market participants and trading tasks are heterogeneous (Müller, Dacorogna, Davé, Pictet, Olsen, and Ward, 1993). In particular, forecasting tasks may involve different trading horizons and multi-period forecast densities.

I suggest that the correlated response model holds in situations which simulate financial situations in addition to the previously studied scenarios. I conjecture that when people are asked to make forecasts from graphically presented price series, the dispersion and variability of the forecasts are correlated with the variability of the data series. Thus, I suggest the following hypotheses:

H<sub>1</sub>: Point forecast dispersion is positively correlated with the local steepness and the oscillation of the data graphs.

H<sub>2</sub>: The local steepness and the oscillation of multi-period forecasts are positively correlated with the corresponding variability measures of the data graphs.

In Experiment 1, Hypothesis H<sub>1</sub> is tested. In particular, I examine its robustness across a wide range of data series, different forecast horizons and when participants are given the option to scale the given data. In Experiment 2, Hypothesis H<sub>2</sub> is tested. I investigate its robustness across a wide range of data series, different forecast densities and when participants are given the option to apply moving average filters on the presented series.

Experiments 1 and 2 are described in Sections 2 and 3. The general discussion is presented in Section 4.

## **2. Experiment 1**

In Experiment 1, participants were presented with a sequence of graphs representing price series. Participants could control the time interval of the graph by using a slider. Similar scaling options are available in financial data analysis programs. Participants were asked to choose the time interval they considered the most appropriate for making financial forecasts and decisions, and then to make

them based on the time-scaled graph. Two variables were manipulated: the variability of the original data graphs, and the required forecast horizon. Figure 1 depicts the task window of Experiment 1.

---

Figure 1 about here

---

## *2.1 Method*

### *2.1.1. Participants*

Thirty-four people (15 men and 19 women) with an average age of 23.29 years acted as participants. Participants' decisions were used to determine their fees: they were paid a flat fee of £3.00 and a further £1.00 if their financial decisions were more than 65% correct. Correctness was determined by participants' performance with respect to the generated graphs. For instance, if prices at the required forecast day were higher than the price on day 200 by more than 5%, a 'buy' decision was considered correct and both 'sell' and 'hold' decisions were considered wrong.

### *2.1.2. Stimulus materials*

The time series presented to the participants were fractional Brownian motions (fBm). Fractal graphs are considered by many researchers adequate for modeling price series, as they attribute more realistic probabilities to financial crises than the random walk model does (Mandelbrot and Hudson, 2004; Mihajlovsky, 2013; Parthasarathy, 2013; Malavoglia, Gaio, Júnior and Lima, 2012; Panas and Ninni, 2010; Sun, Rachev and Fabozzi, 2007; In and Kim, 2006). This is due to the 'fat tails' of a large set of fBm series. Fat tails assign to rare events higher probabilities than those attributed by the normal distribution (Mandelbrot and Hudson, 2004).

Fractional Brownian motions depend on a constant termed the Hurst exponent ( $H$ ). The range of the Hurst exponent of time series is  $[0,1]$ , where the value  $H=0.5$  corresponds to the random walk model

(Peitgen and Saupe, 1988). In this sense, the fractal model can be considered a generalization of the random walk model. However, it was shown that the Hurst exponent of most stocks is in the interval [0.3,0.7] (Sang, Ma, and Wang, 2001).

Examples of fBm series are presented in Figure 2. As can be seen in Figure 2, time series with low Hurst exponent values look noisier and more volatile than graphs with high Hurst exponents. Indeed, the Hurst exponent is negatively correlated with variability measures of the series, including the local steepness, oscillation, and standard deviation (provided that the series were generated by the same algorithm). Gilden, Schmuckler and Clayton (1993) and Kumar, Zhou, and Glaser (1993) showed that people are highly sensitive to the Hurst exponent of fBm graphs.

---

Figure 2 about here

---

In line with the results of Sang, Ma, and Wang (2001), stimulus graphs comprised five sets of three time series with Hurst exponents  $H = 0.3, 0.5,$  and  $0.7$ . Time series were produced using the Spectral method described by Saupe (Peitgen and Saupe, 1988). All series included 62831 ( $\sim 1000 \cdot 2\pi$ ) points. This large number of points was chosen to enable scaling and examination of forecast quality. However, participants were not presented with all data points. The series were presented to the participants as asset price graphs. A constant was added to them to ensure that they were positive. To increase measurement precision by encouraging participants to make forecasts using more than one significant digit, they were also multiplied by 100.

### *2.1.3. Stimulus presentation and control*

Stimulus graphs were presented using a Matlab program that enabled participants to scale the data along the time axis, to make forecasts for a specified horizon, and to express their financial decisions (see Figure 1). Examples of scaled fBm graphs are presented in Figure 2.

Time scaling was accomplished using a slider. At the beginning of each trial, each graph was presented on the time interval [100, 200]. The scaling slider's range varied from a time interval of four days at the maximal zoom-in side of the slider (presentation of price data from days 196 to 200) to 200 days at the maximal zoom-out side of the slider (presentation of price data from days 0 to 200). Thus they could scale the graphs by a factor of 50 (i.e.  $200/4$ ).

Participants made single point forecasts by entering a number into a text box. Forecast horizon was set to 2, 15, or 100 days, making the factor by which horizons varied (i.e.  $100 / 2$ ) identical to that by which scaling could vary (i.e.  $200 / 4$ ).

Participants then made a financial decision to buy another share of the presented asset, to sell their share, or to do neither of these. On each trial, they could change the time interval shown on the graph until they clicked the button "When you are ready, please press OK". They could edit their forecasts until they clicked the button "Save forecast".

### *2.1.4. Design*

Participants were presented with 48 graphs: three familiarization graphs and 45 experimental graphs. Only experimental graphs were included in the analysis. Each graph required three responses: the first was choice of time interval; the second was to forecast the asset's future price; the third was to make a financial decision.

Each participant saw all 15 graphs. Each one was presented three times in different contexts that varied according to the required forecast horizon (2, 15, and 100 days). The order of the graphs and

the required forecast horizons were randomly chosen. This combination produced a three (forecast horizons) by three (Hurst exponent values) by five (instances of time series with the same Hurst exponent values) within-participants design.

#### *2.1.5. Procedure*

Participants were instructed to assume that the experiment day was day 200 and asked to read the following instructions:

“In the following experiment, you are asked to imagine that you are a financial analyst. You have 45 clients. Each of your clients has one share of a single asset. Clients differ in their trading frequency: some clients trade every two days, some trade every 15 days, and some every 100 days. Your aim should be to increase the total value of their portfolios as your fees will depend on your performance.

In order to make your decisions, you will be presented with the price graphs of each of these assets. You will be able to control the time range of each graph by changing its zoom.

For each asset you will be asked to:

1. Notice the trading frequency of your client and the day you will be asked to make financial forecast for. Look at the price graph of the asset carefully.
2. Choose for each graph a time range which you consider the most appropriate for the purpose of making a financial forecast.
3. Write your forecast for the price of the asset on the required day.
4. Advise to your clients whether to buy another share of the asset, sell their share, or hold it.”

Participants could choose the time range of the data graphs by dragging a slider. Forecasts were made by entering a number to a text box. Participants could advise their clients whether to buy, sell,

or hold their shares by clicking one of three buttons. All tasks had to be completed before participants could continue to the next graph.

## 2.2. Results

I excluded from the analysis participants whose means of choices of time scaling factor were more than three standard deviations greater than that of the average for the rest of the group and those whose forecasts were different from the mean of the group by more than two standard deviations. This reduced the size of the sample from 34 to 30 participants, leaving a total of 1350 graphs for the analysis.

Variables of primary interest were the dispersion of participants' forecasts and the local steepness and oscillation of the original graphs. In addition, to assess the extent to which participants used the scaling option and the effect of the scaling on the graphs, the chosen time scaling factor and the local steepness and oscillation of the scaled graphs were extracted.

### 2.2.1. Forecast dispersion

For each combination of Hurst exponent value and forecast horizon, participants were presented with five different graphs (instances). That was taken into account when calculating the forecast dispersion by incorporating different reference level.

Denote by  $H = [0.3, 0.5, 0.7]$  the Hurst exponents of the presented graphs, by  $Instance = [1, 2, 3, 4, 5]$  the graph set labels (in each set, all graphs had the same Hurst exponent value), and by  $Horizon = [2, 15, 100]$  the required forecast horizons.  $F_{i,j,k,l}$  denotes participant  $i$ 's forecast in the condition in which the Hurst exponent of the graph was  $H_j$ , the graph was chosen from set  $Instance_k$ , and the participant was asked to make a forecast for the horizon  $Horizon_l$ . The ranges of these indices are:  $i = 1, \dots, 30, j = 1, 2, 3, k = 1, \dots, 5, l = 1, 2, 3$ . Denote by  $SeriesValue\ onDay200_{j,k}$  the last point of the data series with  $H = H_j$  from set  $k$  which participants were presented with (the value of the series on day 200), and by

$SeriesValueOnForecastDay_{j,k,l}$  the value of the simulated series on the required forecast day.

When an index is omitted, the measure is averaged over the corresponding variable. The following variables were extracted:

1.  $D1_{i,j,k,l} = F_{i,j,k,l} - F_{j,k,l}$ .  $D1$  was used to calculate the forecast dispersion with respect to the mean of the forecasts.
2.  $D2_{i,j,k,l} = F_{i,j,k,l} - SeriesValue onDay200_{j,k}$ .  $D2$  was used to calculate the forecast dispersion with respect to the present price of each asset.
3.  $D3_{i,j,k,l} = F_{i,j,k,l} - SeriesValueOnForecastDay_{j,k,l}$ .  $D3$  was used to calculate participants' forecast error with respect to the produced time series.

Direct forecast dispersion measures are given by the standard deviations of  $D1$ ,  $D2$ , and  $D3$ . The standard deviations of each of these measures in each of the experimental conditions are given in Table 1. As the table shows, when the forecast horizon was 15 days, the standard deviation of  $D1$  decreased when the Hurst exponent increased. For  $D2$  and  $D3$ , similar patterns were found for forecast horizons of two and 15 days.

---

Table 1 about here

---

To examine Hypothesis  $H_1$ ,  $FD1 = |D1|$ ,  $FD2 = |D2|$  and  $FD3 = |D3|$  served as forecast dispersion measures. In addition, the local steepness and oscillation of the original graphs (before scaling) were calculated. To measure the perceived local steepness of a scaled time series, I extracted the average of the absolute value of the gradient at each series point. Then, I multiplied this value by the ratio of the observed time interval and the number of pixels along the time axes of the graph (600) (local steepness measures for the data series after participants' scaling were calculated similarly). The oscillation of each graph was the difference between the graph's minimum and the maximum.

The correlations between  $FD1$ ,  $FD2$  and  $FD3$  and local steepness and oscillation of the original and scaled graphs were significant. These results support Hypothesis  $H_1$ . The correlations are summarized in Table 2. Figure 3 depicts the means of  $FD1_{j,k} = |D1_{j,k}|$  for the different experimental conditions.

---

Table 2 about here

---

---

Figure 3 about here

---

In addition, I carried out for each of the dispersion measures a three-way repeated measures ANOVA using the variables Horizon, Hurst exponent, and ‘instance’ as within-participant variables. I report here the results of the analysis of  $FD1$ . The results of the analysis of  $FD2$  and  $FD3$  were similar.

For  $FD1$ , sphericity assumption was violated for all variables apart from the Hurst exponent and the ‘instance’. The analysis revealed that  $FD1$  was larger when the Hurst exponent was smaller ( $F(2, 58) = 10.32$ ;  $p < .001$ ; partial  $\eta^2 = .26$ ), supporting Hypothesis  $H_1$  (the Hurst exponent is negatively correlated with variability measures).  $FD1$  was larger when forecast horizon was longer ( $F(1.39, 40.42) = 84.67$ ;  $p < .001$ ; partial  $\eta^2 = .75$ ). There was also a significant effect of ‘instance’ on forecast dispersion, indicating that participants reacted to graph characteristics other than the Hurst exponent, too ( $F(4, 116) = 16.91$ ;  $p < .001$ ; partial  $\eta^2 = .37$ ).

All possible interactions between these variables were significant, with  $F > 5.44$  ( $p \leq .002$ ; partial  $\eta^2 > 0.16$ ). I report the results of the interactions and the corresponding simple tests in Table A in the Appendix. None of these interactions contradicted  $H_1$ .

### 2.2.3. Choice of time-scaling factor and properties of the scaled graphs.

To examine whether participants used the scaling option and the effect of the scaling on the graphs, I analyzed the time scaling factors participants chose, as well as the local steepness and oscillation of the scaled graphs.

I refer to the location on the scaling-slider which participants chose for each graph as the *time-scaling factor*. This measurement could vary between 0, corresponding to four days and 1, corresponding to 200 days (the transformation used to translate time-scaling factors to the actual day number presented on the graphs was:  $\text{day number} = 196 * (\text{time-scaling factor}) + 4$ ). The mean time-scaling participants chose was 0.40, and the standard deviation was 0.37.

T-tests performed on participants' choices of scaling factors showed that the mean value was significantly different from 0.5 (the initial setting):  $t(1349) = 9.74, p < .001$ , from 0.0 (maximal zoom-in):  $t(1349) = 40.05, p < .001$ , and from 1.0 (using information from the maximal available time-interval):  $t(1349) = 59.53, p < .001$ . Therefore, it can be concluded that participants used the scaling option.

The local steepness of the scaled graphs was significantly correlated with the local steepness of the original graphs ( $r = .58; p < .01$ ). Similarly, the oscillation of the scaled graphs was significantly correlated with the oscillation of the original graphs ( $r = .58; p < .01$ ). These results show that participants scaled the graphs in a way that preserved the original variability of the graphs.

Figure 4 depicts the mean local steepness and oscillation of the time-scaled graphs for the different conditions of the Hurst exponent and the forecast horizon. To conclude, variability measures of the scaled graphs were correlated with the corresponding variability measures of the data graphs.

---

Figure 4 about here

---

*2.2.4. Additional analysis: A comparison between human forecast dispersion and linear regression prediction intervals for an individual value.*

As a benchmark for the assessment of participants' forecast dispersion, the 95% prediction interval for an individual value of a linear regression (*PI*) was calculated. The length of a 95% *PI* of a linear regression is given by

$$PI = 2 \cdot 1.96 S_e \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{(N - 1)S_x^2}},$$

where  $N$  is the number of data series points,  $\bar{x}$  is the mean of these points,  $S_x$  is the standard deviation of the points, and  $x$  is the forecast point (taking into account the forecast horizon).  $S_e$  is given by  $S_e = \sqrt{\frac{1}{N-2} \sum_{i=1}^N e_i^2}$ , where  $e_i$  is the regression line error, calculated for all data series points (Hyndman and Athanasopoulos, 2014). Here, *PI* was calculated with respect to the time interval participants were initially presented with (the period between day 100 and 200), taking into account  $N = 15,000$  points. The prediction interval lengths, averaged over all graphs presented in each condition, are presented in Table 1.

In order to compare participants' forecast dispersion with *PI*, the standard deviation of *D1* was multiplied by a factor of  $2 \cdot 1.96$ . This is because, assuming normality, the measure  $2 \cdot 1.96 \cdot std(D1)$  represents a 95% confidence interval for *D1*. The values of  $2 \cdot 1.96 \cdot std(D1)$  are presented in Table 1. A comparison of *PI* to  $2 \cdot 1.96 \cdot std(D1)$  reveals that the 95% prediction interval for an individual value of a linear regression is smaller than the human dispersion measure  $2 \cdot 1.96 \cdot std(D1)$  (the same holds also for  $2 \cdot 1.96 \cdot std(D2)$  and  $2 \cdot 1.96 \cdot std(D3)$ ). Thus,

participants' forecast dispersion is larger than linear regression prediction intervals for an individual value. Hence, human forecast dispersion is not statistically optimal, as could be expected from rational forecasters.

### **3. Experiment 2**

In Experiment 2, participants were presented with a sequence of time series. Each sequence was shown on a separate trial. At the beginning of each trial, two identical copies of the same time series were presented on the same axes. Both copies remained visible during the whole duration of each trial. However, the task window enabled participants to smooth one of the graphs. The other graph remained fixed. That made it possible for the participants to smooth each price data graph while seeing the original data. Participants were asked to choose the smoothness level they considered the most appropriate for making financial decisions from it, and then to make forecasts based on the smoothed graph. Two variables were manipulated: the Hurst exponent of the original data graphs (and thus also their local steepness and oscillation), and the number of required forecast points, or, equivalently, the forecast density. Figure 5 depicts the task window of Experiment 2. It shows a graph of the original data and the corresponding smoothed graph (on the same axis).

---

Figure 5 about here

---

#### *3.1. Method*

##### *3.1.1. Participants*

Thirty-four people (15 men and 19 women) with an average age of 26.4 years acted as participants.

They were paid a flat fee of £3.00.

### 3.1.2. Stimulus materials and the presentation program

Stimulus graphs included six sets of five fBm time series with Hurst exponents  $H = 0.3, 0.4, 0.5, 0.6,$  and  $0.7$ . As in Experiment 1, the time series were produced using the Spectral method described by Saupe (Peitgen and Saupe, 1988). All of the time series included 3600 points and were presented to participants as asset price graphs.

Stimulus graphs were presented using a Matlab program. The experimental program enabled participants to apply an averaging filter to the price graphs, while viewing the original price graphs and to make forecasts on pre-specified dates (see Figure 5). Examples of smoothed fBm series are presented in Figure 6.

---

Figure 6 about here

---

Application of the averaging filter was done using a slider. The filter's range was from an averaging window of size 2 (averaging over every two adjacent elements of the series) to averaging over the whole series, the latter resulting in a constant line. To enable participants to both express fine details at the lower end of the scale and reach the maximum averaging, the slider was exponentially calibrated.

The experimental program required participants to make forecasts on dates designated by vertical lines. There were 6, 12, 24, or 36 lines. In each task, participants could change smoothing level until they ticked the box "Completed choice of smoothing level?". They could edit their forecasts by clicking the mouse again on any bar, until they ticked the box "Completed your forecast?".

### 3.1.3. Design

Participants were presented with 23 graphs: three familiarization graphs and 20 experimental graphs. Only experimental graphs were taken into account during the analysis stage. Each graph required two responses. The first response was a choice of smoothing level. The second response was to forecast the asset's future prices.

For each participant, four graphs with each value of Hurst exponent ( $H=0.3, 0.4, 0.5, 0.6, 0.7$ ) were randomly chosen from the stimulus sets. For each value of Hurst exponent, the density of the required forecast was manipulated, and was set to a value of 6, 12, 24, or 36 forecasts within a three-year period. That gave rise to a five (Hurst exponent) by four (forecast density) design. Ordering of trials with different Hurst exponents and forecast densities was random.

*Procedure* Participants were asked to read the following instructions:

“In the following task, you are asked to imagine that you are a financial analyst working at an investment company. Your clients ask you to give them a three year forecast. Each client asks for a forecast of a different resolution: some clients need a monthly forecast (a total of 36 points), some require a forecast point every 6 months (a total of 6 points), and some are interested in an intermediate number of forecast points (a total of 12 points or 24 points).

You will be presented with a series of 3 practice graphs and 20 experiment graphs representing prices of different assets. The program will enable you to set the smoothness level of the data graphs. You are asked:

1. to look at the graphs carefully,
2. for each of the graphs, to determine the smoothness level you consider the most appropriate for making financial decisions from it,

3. to predict the prices on a series of time points based on the smoothed graph. The number of forecasts will be 6, 12, 24, or 36 points according to the request obtained from each of your clients.”

Participants chose a smoothness level of data graphs by dragging a slider. The smoothed graph was presented in red. The original graph was presented in blue.

Forecasts were made by clicking a mouse at specific dates, designated by vertical lines. Participants had to complete the forecasts on all vertical lines (dates) before they could continue to the next graph.

### *3.2. Results*

Participants whose means of smoothing level choices were more than two standard deviations greater than that of the average for the rest of the group were excluded from the analysis. This reduced the size of the sample from 34 to 32 participants. Three additional extreme measurements (out of the original  $20 * 34 = 680$  measurements), in which participants chose smoothing levels more than four standard deviations greater than that of the mean of the experimental condition were removed as well. Therefore, 637 graphs were used for the analysis.

The variables of primary interest were the local steepness and oscillation of the forecasts and the original data graphs. The standard deviation of participants' forecasts was extracted to provide comparison of the results with an additional variability measure. As in Experiment 1, the local steepness and oscillation of smoothed graphs were calculated. The chosen smoothing factors were extracted, too, as they could indicate the window sizes of the moving average filters which participants applied on the data. The local steepness and the oscillation of smoothed graphs were used to assess the similarity between the original and smoothed data.

### 3.2.1. Properties of participants' forecasts

To examine Hypothesis H<sub>2</sub>, the local steepness and oscillation of the forecasts and the original data graphs were extracted. Significant correlations were found between the local steepness of the forecasts and the local steepness of the data before and after the smoothing ( $r = 0.39$ ;  $p < .01$ , and  $r = .33$ ;  $p < .01$  respectively). Significant correlations were found between the oscillation of the forecasts and the oscillation of the data both before and after smoothing ( $r = .43$ ;  $p < .01$ , and  $r = .40$ ;  $p < .01$  respectively). Positive correlations were found between the standard deviation of the forecasts and the local steepness and oscillation of the data graphs before smoothing ( $r = 0.375$ ;  $p < .01$ , and  $r = .443$ ;  $p < .01$  respectively). These results support Hypotheses H<sub>2</sub>. They suggest that participants imitate properties of the data when making judgmental forecasts from graphs, hence provide further support for the correlated response model. In addition, they are in line with Harvey's finding, that people tend to add noise to their forecasts (1995). This tendency is irrational in the sense that trend-line statistical forecasts are characterized by superior accuracy (Harvey, Ewart, and West, 1997).

To obtain further support for H<sub>2</sub>, I carried out a two-way repeated measures ANOVA on the standard deviation of participants' forecasts using the Hurst exponent and forecast density as within-participant variables. Huynh-Feldt test showed that the standard deviation of participants' forecasts was larger when Hurst exponent of the data graphs was smaller ( $F(3.58, 110.97) = 32.95$ ;  $p < .01$ ; partial  $\eta^2 = .52$ ). (The effect of the forecast density on forecast dispersion was insignificant,  $p = 0.057$ ). Figure 7 depicts the standard deviation of participants' forecasts against the required forecast density for the given values of the Hurst exponent.

---

Figure 7 about here

---

### 3.2.2. Choice of smoothness level and properties of the smoothed data graphs

The mean smoothness level participants chose was 59.09. The standard deviation was larger than the mean: 82.61. A t-test performed on participants' choices of smoothness levels showed that it was significantly larger than 1 (a trivial filter):  $t(636) = 17.76$  ( $p < .01$ ).

The correlation between the oscillations of the smoothed and original data graphs was  $r = .88$ ;  $p < .01$ . That suggests that participants smoothed the graphs in a way that maintained a correlation between variability measures of the smoothed graphs and the original graphs.

Figure 8 depicts the local steepness and mean oscillation of the smoothed data graph for the different values of the Hurst exponent and the different numbers of required forecast points.

---

Figure 8 about here

---

## 4. General discussion

In the book "An Engine, Not a Camera: How Financial Models Shape Markets", MacKenzie (2006, page 12) wrote: "Financial economics, I argue, did more than analyze markets; it altered them. It was an "engine" [...]: an active force transforming its environment, not a camera passively recording it". I suggest that a central component of the engine of the market – the way people perceive financial data and make forecasts from it - is not as passive as a camera; yet, it preserves important properties of the data.

The way people perceive graphically presented price series was examined through their choices of scaling and moving window averaging. Both techniques are highly popular among investors and traders (Batchelor, 2013; Batchelor and Kwan, 2007; Cheung and Chinn, 2001; Gehrig and Menkhoff, 2006; Menkhoff, 2010; Taylor and Allen, 1992). The results of this study indicate that, though there is a large variability among people in the choice of scaling and moving window averaging parameters, there are significant correlations between the local steepness and oscillation of the transformed data graphs, and the local steepness and oscillation of the original price graphs.

Two properties of forecasts were examined: forecast variability and forecast dispersion. Significant correlations were found between both of these measures and the local steepness and oscillation of the original price graphs. Furthermore, the Hurst exponent had a significant effect on participants' forecasts variability and dispersion. These results were in line with previously reported findings (Reimers and Harvey, 2011; Harvey, 1995; Bolger and Harvey, 1993; Harvey, Ewart, and West, 1997), and were robust across a wide range of data series, different tasks (point and multi-period forecasts) and when participants could scale or smooth the data. The robustness of these findings with respect to scaling and smoothing can be partially explained by the correlations between variability measures of the scaled and smoothed graphs and the original data graphs.

#### *4.1. Contribution*

This paper contributed to two fields: Finance and Judgmental Forecasting. As forecast dispersion was shown to be a predictor of important financial variables (Li and Wu, 2014; Athanassakos and Kalimipalli, 2003) many studies focused on the determinants of forecast dispersion (e.g. Platikanova and Mattei, 2016; Kwon, 2002). However, the effect of the properties of the data on forecast dispersion was overlooked. Given that a large percentage of market participants incorporates in decision making processes financial graph analysis (Menkhoff, 2010; Gehrig and Menkhoff, 2006), it was important to examine the effect of price graph properties on forecast dispersion. The results

presented here indicate that forecast dispersion is significantly correlated with historical price variability measures.

Research in Judgmental Forecasting tended to focus on the identification of specific biases and the estimation of forecast errors. Uniting the results of a few classical papers (Reimers and Harvey, 2011; Harvey, 1995; Bolger and Harvey, 1993; Harvey, Ewart, and West, 1997), I suggested the correlated response model, according to which, in judgmental forecasting tasks from graphically presented time series, people tend to imitate the given data. This tendency may be perceived as irrational, in the sense that participants included noise in their forecasts. In addition, human point forecast dispersions were larger than 95% prediction intervals of linear regressions.

Furthermore, the experiments described in this study explored the effects of a few factors which had not been studied before, including scaling and smoothing of data graphs. Aiming at high external validity, complex time series resembling financial price series were utilized as stimuli (Mandelbrot and Hudson, 2004).

#### *4.2. Limitations and further research*

Though price series volatility changes dynamically, it is approximated many times by a slowly varying function (Zhu and Chen, 2011), or a piecewise constant function (Lu, 2010). The results presented here showed that forecast dispersion is correlated with past volatility measures. However, forecast dispersion is likely to be also correlated with volatility of future returns (Athanasakos and Kalimipalli, 2003; Lobo and Tung, 2000). Therefore, judgmental forecasts may serve as a mechanism which helps stabilize price series volatility. A field research exploring this conjecture could yield an important insight into the nature of volatility.

Participants in Experiments 1 and 2 were lay people. Research comparing financial forecasts of lay people and practitioners has typically found only small differences between the two groups

(Zaleskiewicz, 2011; Muradoğlu and Önköl, 1994). Bodnaruk and Simonov (2015) have recently showed that financial experts (fund managers) do not make better investment decisions than other investors. Moreover, during the last years, the internet has made trading easier for lay people (Muradoglu and Harvey, 2012) and inexperienced investors (Barber and Odean, 2001). Nevertheless, lack of expertise in the sample could limit the generality of the results. Therefore, I consider the study of the effects of expertise on performance in the tasks employed here worthwhile.

The stimulus graphs utilized in this study displayed synthetic fractal series. The choice of fractals aimed at a high external validity: though previous studies in judgmental Forecasting usually utilized less complex series as experimental stimuli, fractals were employed in many financial studies and especially those which dealt with trading horizons (e.g. Kristoufek, 2012). However, it is important to note that not all researchers accept the fractal market hypothesis (Onali and Goddard, 2011). Therefore, it is important to replicate the experiments performed here using real-life price series. It could be also beneficial to examine the effect of the graphical interface on participants' responses. The graphical interfaces developed for the experiments may limit the generality of the results.

Furthermore, the graphs that participants were presented with did not depict special financial situations, such as financial crises and bubbles. I consider it essential to explore participant behavior in such cases.

Finally, in Experiment 1, trading horizon and the Hurst exponents of the graphs were treated as independent variables. However, this assumption may not hold in all conditions (Vácha and Vošvrda, 2005).

## **Acknowledgement**

I would like to thank Prof. Nigel Harvey for his help, guidance, and encouragement.

I would like to thank the reviewers for valuable comments.

## **Appendix: Interactions and tests of simple effects**

---

Table A about here

---

## References

- Athanassakos, G. & Kalimipalli, M. (2003). Analyst forecast dispersion and future stock return volatility. *Quarterly Journal of Business and Economics*, 42 (1/2), 57-78. ISSN: 07475535.
- Bandura, A., & Barab, P. G. (1971). Conditions governing nonreinforced imitation. *Developmental Psychology*, 5(2), 244-255. doi: 10.1037/h0031499.
- Barber, B. M., & Odean, T. (2001). The Internet and the investor. *Journal of Economic Perspectives*, 15(1), 41-54. doi: 10.1257/jep.15.1.41.
- Batchelor, R. (2013). Forecasting financial markets: Some light from the dark side. Forecasting with big data, The 33<sup>rd</sup> International Symposium on Forecasting, Proceedings.
- Batchelor, R., & Kwan, T. Y. (2007). Judgemental bootstrapping of technical traders in the bond market. *International Journal of Forecasting*, 23, 427–445. doi: 10.1016/j.ijforecast.2007.05.007.
- Bodnaruk, A., & Simonov, A. (2015). Do financial experts make better investment decisions? *Journal of Financial Intermediation*, 24, 514-536. doi: 10.1016/j.jfi.2014.09.001.
- Bolger, F., & Harvey, N. (1993). Context-sensitive heuristics in statistical reasoning. *The Quarterly Journal of Experimental Psychology*, 46 (4), 779–811. doi: 10.1080/14640749308401039.
- Caspi, A., & Moffitt, T. E. (1993). When do individual differences matter? A paradoxical theory of personality coherence. *Psychological Inquiry*, 4(4), 247-271. doi: 10.1207/s15327965pli0404\_1.
- Cheung, Y. W., & Chinn, M. D. (2001). Currency traders and exchange rate dynamics: a survey of the US market. *Journal of International Money and Finance* 20, 439–471. doi: 10.1016/S0261-5606(01)00002-X.

- Cukierman, A., & Wachtel, P. (1979). Differential inflationary expectations and the variability of the rate of inflation: Theory and evidence. *The American Economic Review*, *69*(4), 595-609. ISSN: 00028282.
- Cukierman, A., & Wachtel, P. (1982). Relative price variability and nonuniform inflationary expectations. *Journal of Political Economy*, *90*(1), 146-157. doi: 10.1086/261043.
- Gehrig, T., & Menkhoff, L. (2006). Extended evidence on the use of technical analysis in foreign exchange. *International Journal of Finance & Economics*, *11*(4), 327-338. doi: 10.1002/ijfe.301.
- Gilden, D. L., Schmuckler, M. A., & Clayton, K. (1993). The perception of natural contour. *Psychological Review*, *100* (3), 460-478. doi: 10.1037/0033-295X.100.3.460.
- Glezakos, M., & Mylonas, P. (2003). Technical analysis seems to be a valuable investment tool in the Athens and Frankfurt stock exchanges. *European Research Studies*, *6* (1-2), 169-192.
- Harvey, N. (1995). Why are judgments less consistent in less predictable task situations? *Organizational Behavior and Human Decision Processes*, *63* (3), 247–263. doi: 10.1006/obhd.1995.1077.
- Harvey, N., Ewart, T., & West, R. (1997). Effects of data noise on statistical judgement. *Thinking & Reasoning*, *3* (2), 111-132. doi: 10.1080/135467897394383.
- Heyes, C. (2011). Automatic Imitation. *Psychological Bulletin*, *137*(3), 463-483. doi: 10.1037/a0022288.
- Hyndman, R. J., & Athanasopoulos, G. (2014). Forecasting: Principles and practice. OTexts.com. Available on <https://www.otexts.org/fpp>.

- In, F., & Kim, S. (2006). Multiscale hedge ratio between the Australian stock and futures markets: Evidence from wavelet analysis. *Journal of Multinational Financial Management* 16 (4), 411-423. doi:10.1016/j.mulfin.2005.09.002.
- Kristoufek, L. (2012). Fractal markets hypothesis and the global financial crisis: scaling, investment horizons and liquidity. *Advances in Complex Systems*, 15, 1250065. doi: 10.1142/S0219525912500658.
- Kumar, T., Zhou, P., & Glaser, D. A. (1993). Comparison of human performance with algorithms for estimating fractal dimension of fractional Brownian statistics. *Journal of the Optical Society of America. A, Optics and image science*, 10 (6), 1136-1146. ISSN: 0740-3232.
- Kwon, S. (2002). Financial Analysts' Forecast Accuracy and Dispersion: High-Tech versus Low-Tech Stocks. *Review of Quantitative Finance and Accounting*, 19(1), 65-91 doi: 10.1023/A:1015730325706.
- Lawrence, M., Goodwin, P., O'Connor, M., & Önköl, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22 (3), 493-518. doi: 10.1016/j.ijforecast.2006.03.007.
- Lawrence, M., & Makridakis, S. (1989). Factors affecting judgemental forecasts and confidence intervals. *Organizational Behavior and Human Decision Processes*, 43 (2), 172–187. doi: 10.1016/0749-5978(89)90049-6.
- Li, M. Y., & Wu, J. S.(2014). Analysts' Forecast Dispersion and Stock Returns: A Quantile Regression Approach. *Journal of Behavioral Finance*, 15(3), 175-183 doi: 10.1080/15427560.2014.942420.

- Lobo, G., & Tung, S. (2000). Financial analysts' earnings forecast dispersion and intraday stock price variability around quarterly earnings announcements. *Review of Quantitative Finance and Accounting*, 15(2), 137-151. doi: 10.1023/A:1008317129991.
- Lu, B. (2010). Recovering a piecewise constant volatility from perpetual put option prices. *Journal of Applied Probability*, 47(3), 680-692. ISSN: 00219002.
- MacKenzie, D. (2006). *An engine, not a camera: How financial models shape markets*. Massachusetts, USA: The MIT Press.
- Malavoglia, R. C., Gaio, L. G., Júnior, T. P., & Lima, F. G. (2012). The Hurst exponent: a study of the major international stock markets. *Journal of International Finance & Economics*, 12 (1), 113-121. ISSN: 1555-6336.
- Mandelbrot, B., & Hudson R. L. (2004). *The (mis)behaviour of markets, a fractal view of risk, ruin and reward*. London, UK: Profile Books LTD.
- Menkhoff, L. (2010). The use of technical analysis by fund managers: International evidence. *Journal of Banking and Finance*, 34(11), 2573-2586. doi: 10.1016/j.jbankfin.2010.04.014
- Mihajlovsky, O. (2013). Fractal characteristics of world market commodity derivatives. *Annales Universitatis Apulensis : Series Oeconomica*, 15(2), 662-669. ISSN: 14549409.
- Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B., & Ward, J. R. (1993). Fractals and intrinsic time – A challenge to econometricians. 39th International AEA Conference on Real Time Econometrics, 14–15 October 1993, Luxembourg.
- Muradoglu, G., & Harvey, N. (2012). Behavioural finance: the role of psychological factors in financial decisions. *Review of Behavioral Finance*, 4 (2), 68 – 80. doi: 10.1108/19405971211284862.

- Muradođlu, G. & Önkai, D. (1994). An exploratory analysis of portfolio managers' probabilistic forecasts of stock prices. *Journal of Forecasting*, 13, 565–578. Available at SSRN: <http://ssrn.com/abstract=1296864>.
- Onali, E., & Goddard, J. (2011). Are European equity markets efficient? New evidence from fractal analysis. *International Review of Financial Analysis*, 20(2), 59-67. doi: 10.1016/j.irfa.2011.02.004.
- Panas, E., & Ninni, V. (2010). The distribution of London Metal Exchange prices: A test of the fractal market hypothesis. *European Research Studies*, 13(2), 193-210.
- Parthasarathy, S. (2013). Long range dependence and market efficiency: Evidence from the Indian stock market. *Indian Journal of Finance*, 7 (1), 17-25.
- Peitgen, H. O., & Saupe, D. (1988). *The science of fractal images*. New York, NY: Springer-Verlag.
- Platikanova, P., & Mattei, M. (2016). Firm geographic dispersion and financial analysts' forecasts. *Journal of Banking & Finance*, 64, 71. ISSN: 03784266.
- Reimers, S., & Harvey, N. (2011). Sensitivity to autocorrelation in [judgmental time series](#) forecasting. *International Journal of Forecasting*, 27(4), 1196-1214. doi: 10.1016/j.ijforecast.2010.08.004.
- Sang, H-w., Ma, T., & Wang, S-z. (2001). Hurst exponent analysis of financial time series. *Journal of Shanghai University (English Edition)*, 5 (4), 269-272. ISSN: 1007-6417.
- Sun, W., Rachev, S., & Fabozzi, F. J. (2007). Fractals or I.I.D.: Evidence of long-range dependence and heavy tailedness from modeling German equity market returns. *Journal of Economics and Business*, 59 (6), 575-95. doi:10.1016/j.jeconbus.2007.02.001.

Taylor, M. P., & Allen, H. (1992). The use of technical analysis in the foreign exchange market.

*Journal of International Money and Finance*, 11 (3), 304–314. doi: 10.1016/0261-5606(92)90048-3.

Trench, W. F. (2002). Introduction to real analysis. Prentice Hall.

([http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH\\_REAL\\_ANALYSIS.PDF](http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF)).

Vácha, L. & Vošvrda, M. S. (2005). Dynamical agents' strategies and the fractal market hypothesis.

*Prague Economic Papers*, 14(2), 163-70.

Yang, H. I. (2012). Capital market consequences of managers' voluntary disclosure styles. *Journal of*

*Accounting and Economics*, 53(1-2), 167-184. DOI: 10.1016/j.jacceco.2011.08.003.

Zaleskiewicz, T. (2011). Financial forecasts during the crisis: Were experts more accurate than

laypeople? *Journal of Economic Psychology*, 32 (3), 384–390. doi:

10.1016/j.joep.2011.02.003.

Zhu, S. P., & Chen, W. T. (2011). Should an American option be exercised earlier or later if volatility is

not assumed to be a constant? *International Journal of Theoretical and Applied Finance*,

14(08), 1279-1297. doi: 10.1142/S0219024911006851.