EXPECTATIONS, LEARNING GAINS, AND FORECAST ERRORS:
ASSESSING NONLINEARITIES WITH A FUNCTIONAL COEFFICIENT APPROACH

FABIO MILANI

University of California, Irvine

Abstract. This paper investigates potential nonlinearities in the gain function, which, under adaptive learning, regulates the updating of agents' beliefs in response to recent forecast errors.

I use data on professional survey forecasts to estimate nonparametric functional-coefficient regression models.

The estimation results reveal nonlinearities in the relationships between expectations and forecast errors, which are indicative of nonlinear gain functions. Gains increase when forecast errors are historically large, and respond asymmetrically to past overpredictions and underpredictions. The findings suggest incorporating nonlinearities in the modeling of learning gains, instead of relying on the constant-gain assumption.

Keywords: Survey Forecasts, Nonlinear Gain, Adaptive Learning, Nonparametric Regression, Functional Coefficient Regression Model.

JEL classification: C14, E31, E32, E70.

Date: July, 2025.

Addresses for Correspondence: Department of Economics, 3151 Social Science Plaza, University of California, Irvine, CA 92697-5100, U.S.A. Email: fmilani@uci.edu. Homepage: https://sites.socsci.uci.edu/∼fmilani.

1. Introduction

Expectations play a central role in macroeconomics, as they affect decisions by households, firms, and policymakers. For decades, macroeconomic models almost universally relied on the rational expectations hypothesis (REH). But, over time, a growing body of literature has documented and incorporated departures from rational expectations, with the adaptive learning approach (e.g., Evans and Honkapohja, 2001) emerging as a leading alternative to the REH.

Under adaptive learning, agents are often assumed to revise beliefs according to a constant-gain algorithm.[1] Beliefs are revised in the direction of the most recent forecast error, but with an identical learning coefficient each period.

The constant gain assumption may represent a possible limitation of conventional learning models. The learning rate could be a nonlinear function of third variables, rather than a scalar. For example, agents might forecast based on loss functions that diverge from the MSE loss: they may neglect small forecast errors and weigh heavily larger forecast errors; moreover, they may weigh over- and under-predictions asymmetrically. Such loss functions are bound to introduce nonlinearities in the gain function.

This paper uses data on survey expectations about inflation, output, and interest rates, along with the corresponding forecast errors, to assess the importance of nonlinearities in the formation of expectations and in the agents' learning gain. To maintain intuition, I assume a simple learning model, with agents that learn about long-run values of variables (i.e., intercepts in their perceived law of motion), to capture the relationships between expectations and forecast errors, and how they are driven by the learning gain.

In the empirical analysis, I first provide nonparametric evidence using Kernel regressions. Then, I exploit more structure by recognizing that gain functions are likely to depend on past forecast errors and I estimate the relationships between changes in survey expectations and forecast errors using a functional-coefficient regression model based on Cai *et al.* (2000). The approach can be seen as semiparametric, since it imposes more restrictions on the coefficients than a purely nonparametric alternative.

The results support the presence of nonlinearities. In particular for inflation, the gain coefficient is lower when forecast errors are closer to zero and increases for larger, whether

[1] The assumption is common in empirical work: Milani (2007, 2011, 2017), Slobodyan and Wouters (2012a,b), and Eusepi *et al.* (2025).

positive or negative, forecast errors. More uncertainty characterizes the results for output forecasts, but gains still generally increase with the size of forecast errors. The relationship between interest rate expectations and forecast errors is close to linear, and gains roughly constant, except in cases of extreme forecast errors.

The results suggest a revision of learning models to incorporate endogenously changing, nonlinear, gain functions.

2. Expectation Formation and Belief Updating

I assume a simple expectation formation model, according to which agents attempt to learn about intercepts, or long-run values, for inflation, output, and interest rates. These are the variables that agents would need to forecast in a textbook New Keynesian model (e.g., Woodford, 2003). Despite its simplicity, learning about long-run means has been shown to be crucial for matching data on survey forecasts (e.g., Farmer *et al.*, 2024, Patton and Timmermann, 2010), and a similar approach has been recently used in Eusepi *et al.* (2025), among others.

Agents adopt a Perceived Law of Motion (PLM):

$$\begin{bmatrix} \pi_t \\ \Delta y_t \\ i_t \end{bmatrix} = Z_t = \mu_t + \epsilon_t \tag{1}$$

and form expectations (here for $t+1$) as:

$$\widehat{E}_t Z_{t+1} = \widehat{\mu}_t, \tag{2}$$

where $\widehat{E}_t Z_{t+1} = [\widehat{E}_t \pi_{t+1}, \widehat{E}_t \Delta y_{t+1}, \widehat{E}_t i_{t+1}]'$ and $\widehat{\mu}_t = [\widehat{\mu}_t^{\pi}, \widehat{\mu}_t^{\Delta y}, \widehat{\mu}_t^{i}]'$. Beliefs are updated over time following a constant-gain learning algorithm:

$$\widehat{\mu}_t = \widehat{\mu}_{t-1} + \bar{g} \cdot fe_t \tag{3}$$

where $\bar{g}$ denotes the constant-gain coefficient and $fe_t = Z_t - \widehat{E}_{t-1} Z_t = Z_t - \widehat{\mu}_{t-1}$ denotes the most recent forecast error.[2]

[2]Under learning, there can be a second equation reflecting the updating of the precision matrix R_t or, for stochastic-gradient learning: $R_t = I$.

Under this learning model, the change in expectations between two periods is related to the most recent forecast errors as

$$\widehat{E}_t Z_{t+1} - \widehat{E}_{t-1} Z_t = \widehat{\mu}_t - \widehat{\mu}_{t-1} = \bar{g} \cdot fe_t. \tag{4}$$

Therefore, we should expect a clear connection between changes in expectations and recent forecast errors. With a constant gain, the relationship is linear. In the empirical analysis, the paper will investigate potential nonlinearities in this relation, which would be suggestive of a nonlinear gain function $g(\cdot)$ instead.

3. Data

Data on expectations are obtained from the Survey of Professional Forecasters. I use forecasts for GDP Deflator Inflation, Real GDP Growth, and 3-month T-bill interest rates formed for quarter t ("nowcasts") and $t+1$.[3] Inflation and GDP forecasts are available from 1968:Q4 to 2025, whereas forecasts for interest rates are available from 1982:Q4 to 2025.

I compute the relevant forecast errors at each horizon, using the first vintage of release for realized variables.

4. Functional Coefficient Model

4.1. **Nonparametric Evidence.** I start by presenting preliminary evidence on possible nonlinearities between expectations and forecast errors using a nonparametric approach. Let $m(x) = E(Y_t|X_t = x)$ denote the conditional expectation of Y_t given $X_t = x$ and consider a Taylor expansion for x around x_0:

$$m(x) = m(x_0) + m'(x_0)(x - x_0) + ... + \frac{m^{(p)}(x_0)}{p!}(x - x_0)^p + O\{(x - x_0)^{p+1}\}. \tag{5}$$

Around x_0, $m(x)$ can then be modeled statistically as

$$m(x) \approx \sum_{j=0}^{p} \beta_j (x - x_0)^j, \tag{6}$$

[3] I keep the focus on short-term forecasts in this paper; I've also analyzed forecasts at horizons 2 to 4, but their relationships with forecast errors are more tenuous. I've also used different measures of interest-rate expectations, extracted from T-bills at different maturities using the expectations hypothesis of the term structure, and the results are very similar.

where each $\beta_j = m^{(j)(x_0)/j!}$ represents a local parameter. I use a local polynomial Kernel regression estimator, which minimizes the weighted sum of squared residuals

$$m(x) = \sum_{t=1}^{T} \left(Y_t - \sum_{j=0}^{p} \beta_j (X_t - x_0)^j \right)^2 K\left(\frac{X_t - x_0}{h}\right) \tag{7}$$

where K denotes the Kernel function, and h the bandwidth. In the analysis, I use a local linear approximation (i.e., $p = 1$).[4]

Figure 1 shows the estimated relationships using nonparametric Kernel regressions between changes in expectations $\widehat{E}_t Z_{t+h} - \widehat{E}_{t-1} Z_{t-1+h}$ and the most recent forecast errors fe_t.

The top panels show the Kernel estimates for the nowcasts and for one-period-ahead-expectations for inflation. Inflation nowcasts are relatively insensitive to forecast errors. Only very positive or very negative forecast errors cause changes in expectations. One-quarter-ahead inflation expectations, instead, react more significantly to forecast errors. The slope of the relationship remains roughly constant for forecast errors around zero, but it steepens when those are large and positive and, particularly, when they are large and negative.

The third and fourth panels present the Kernel regression estimates for output growth expectations. Again, nowcasts are largely unresponsive, except when forecast errors are largest. The relationship between one-period-ahead expectations and forecast errors has a gentle slope in correspondence of most forecast errors, but there are clear shifts for very large magnitudes, causing more substantial movements in expectations.

The bottom panels show the results for expectations about the 3-month T-bill rate. The relationship between one-quarter-ahead changes in expectations and forecast errors is close to linear; the slope again changes only for larger forecast errors: it steepens for negative errors, but it flattens for positive ones. Same-quarter forecasts display a smaller, but positive, slope (with minimal forecast errors).

4.2. **Functional Coefficient Regressions.** The Kernel estimates in Figure 1 are suggestive of potential nonlinearities in the relations between expectations and forecast errors. In our learning model, nonlinearities would emerge due to deviations of the gain coefficient from the conventional constant-gain assumption. The gain may, in fact, be endogenous and its value shift depending on the magnitude and direction of the most recent forecast errors.

[4]With $p = 0$, we would have instead the Nadaraya-Watson estimator.

The relationship between changes in expectations and forecast errors in (4) can then be re-expressed as

$$\Delta \widehat{E}_t Z_{t+1} = \widehat{\mu}_t - \widehat{\mu}_{t-1} = g(fe_{t-1})fe_t, \tag{8}$$

where the function $g(\cdot)$ now allows for a nonlinear dependence of the gain on past forecast errors. In the learning model, agents enter period t with their updated beliefs from the previous period $t-1$ and with an updated gain coefficient that can vary based on the previous period's forecast errors. They would observe the new forecast errors, revise their beliefs, and form their expectations accordingly.

To investigate the nonlinearity in the expectational data, I use a functional-coefficient regression (FCR) model, based on Cai *et al.* (2000). Compared to a fully-nonparametric approach, functional coefficient modeling allows similar flexibility in fitting the data while avoiding the curse of dimensionality. The additional structure it imposes on the coefficients helps their interpretation. At the same time, the FCR approach remains less restrictive than a parametric nonlinear regression, which requires researchers to specify the form of the nonlinearity *a priori*, among infinitely many possibilities.

The FCR model can be written as

$$Y_t = a_0(u_t) + a_1(u_t)X_{1,t} + ... + a_k(u_t)X_{k,t} + \varepsilon_t, \tag{9}$$

which shows how the coefficients $a(\cdot)$ vary based on the values of a "threshold" variable u_t.

The unknown coefficients in (9) are estimated using local linear regressions. In a neighborhood of u_0, the function $a(\cdot)$ can be approximated using a first-order Taylor expansion:

$$a_j(u) \approx a_j(u_0) + a_j'(u_0)(u-u_0) = a_j + b_j(u-u_0). \tag{10}$$

Then, a local linear regression is run using observations $\{U_i, X_i, Y_i\}_{i=1}^n$, where $n = T - p$ and $X_i = (X_{i1}, ..., X_{ip})^T$, to minimize

$$\sum_{i=1}^{n}\left[Y_i - \sum_{j=1}^{p}(a_j + b_j(U_i - u_0))X_{ij}\right]^2 K\left(\frac{U_i - u_0}{h}\right), \tag{11}$$

with respect to a_j, b_j. The local linear regression estimator is given by $\hat{a}_j(u_0) = \hat{a}_j$, $j = 1, ..., n$, and it can be expressed in Kernel form as

$$\hat{a}_j(u_0) = \sum_{k=1}^{n} K_{n,j}(U_k - u_0, X_k)Y_k, \tag{12}$$

with

$$K_{n,j}(u, x) = e_{j,2p}^T (\tilde{X}^T W \tilde{X})^{-1} \begin{pmatrix} x \\ ux \end{pmatrix} K(\frac{u}{h}), \tag{13}$$

$e_{j,2p}^T$ a selection vector, $\tilde{X} = \left(X_i^T, X_i^T(U_i - u_0)\right)$, and $W = \left\{K\left(\frac{U_1 - u_0}{h}\right), ..., K\left(\frac{U_n - u_0}{h}\right)\right\}$. Following Fan and Yao (2003), I start with a pilot bandwidth, chosen using the Modified Multi-Cross-Validation approach proposed by Cai *et al.* (2000),[5] and then, in each case, select the optimal final bandwidth to minimize the Mean Integrated Square Error (MISE), which accounts for both the squared bias and variance at each evaluation point.[6] In all cases, I employ an Epanechikov Kernel function.[7]

Cai et. al (2000) demonstrate that the FCR procedure can successfully recover common nonlinear functions (such as TAR and EXPAR models) when they are imposed using simulated data under known DGPs.

In the case of our adaptive learning model, the specific FCR equation corresponding to (9) is given by

$$\Delta \widehat{E}_t Z_{t+1} = a_0(fe_{t-1}) + a_1(fe_{t-1})fe_t + \varepsilon_t, \tag{14}$$

which captures the nonlinearity of the gain function and the gain's dependence on recent forecast errors (with fe_{t-1} serving as the threshold variable u_t).

Figure 2 shows the results of the FC regressions, focusing on one-period-ahead expectations, the most relevant case. The panels show the coefficients $a_1(fe_{t-1})$, which specifically

[5]The procedure requires a choice of m and Q, such that $n > mQ$, where n is the number of observations, and where the approach consists of using Q subseries of length $n - qm$, $q = 1, ..., Q$ to estimate the unknown coefficient functions and obtain prediction errors. The pilot bandwidth is then chosen as the one minimizing the average prediction error. I use $m = 0.1 * n$, and Q=4, as suggested in Fan and Yao (2003).

[6]As explained in Fan and Gijbels (1996), the order of the pilot polynomial needs to exceed the estimation order by an even integer (2 in this case).

[7]The optimal bandwidths selected in the estimation are 0.918 for inflation, 0.391 for output, and 0.198 for interest rates.

capture how the sensitivity of changes in expectations to new forecast errors varies as a function of previous forecast errors. The coefficient corresponds to the gain value in (4) or (8): a flat line would suggest a constant gain, whereas more complicated shapes in the functional coefficient would indicate nonlinearities in the gain function.

The estimation results point to non-linearities in the data. The FCR estimates suggest gains that typically increase in the magnitude of forecast errors.

In the case of inflation expectations, the functional coefficient (or gain) is at its lowest (just below 0.1) in correspondence of past forecast error values that are slightly below zero. More extreme forecast errors lead to larger gain coefficients; gains are at their highest when forecast errors are large and negative, suggesting that agents revise their expectations more rapidly when they recently overestimated inflationary pressures.[8]

The results are more mixed for output expectations. If one disregards results in the tails, a similar nonlinearity exists in the data. The gain is lower (with values between 0.07 and 0.13) when forecast errors fall around zero, but it increases with larger forecast errors. The results, however, aren't as clear-cut, since more uncertainty exists on the tails of the forecast error distribution, with coefficients reverting to lower values.

For interest rate expectations, the relationship between expectations and forecast errors is close to linear, and hence the gain function close to constant, for large parts of the forecast error distribution. Gain values are also much larger. But for extreme forecast errors, nonlinearities reappear, with increasing gains after recent interest rates underpredictions and smaller gains after overpredictions.

5. Conclusions

Adaptive learning provides one of the most popular alternatives to the REH in macroeconomics. Learning models assume agents who revise their beliefs in the direction of recent forecast errors, typically with a constant-gain coefficient. They imply identical rates of update every period regardless of the magnitude and sign of forecast errors.

Here, I showed, using nonparametric approaches, that nonlinearities in the gain functions should deserve a bigger role in future research. In particular, learning models should incorporate endogenous gains that increase when forecast errors exceed a certain magnitude,

[8] Gain coefficients are higher here than estimated in DSGE models, since agents learn only about intercepts and not dynamic coefficients.

and that are potentially asymmetric between positive and negative forecast errors; such gain functions would conform more closely to the empirical evidence. Learning with a constant gain remains, however, a reasonably accurate approximation as long as past forecast errors fall in the center region of their historical distribution. The findings closely mirror those in Gáti (2023) for long-run inflation expectations. She estimates an endogenous gain function, which is also nonlinear and asymmetric, and shows that incorporating the nonlinearity leads to important differences in the optimal monetary policy responses.[9]

For brevity, this paper has considered a simplified learning model as a first step. In future research, the agents' perceived model can be extended to match the MSV solution under RE. Other nonlinearities can be studied, including different specifications (e.g., TAR and STAR models) or different threshold variables.

References

[1] Cai, Z., Fan, J., & Yao, Q. (2000). "Functional-Coefficient Regression Models for Nonlinear Time Series". *Journal of the American Statistical Association*, 95(451), 941-956.

[2] Eusepi, S., Giannoni, M., & Preston, B. (2025). "The Short-Run Policy Constraints of Long-Run Expectations". *Journal of Political Economy*, forthcoming.

[3] Evans, G.W., and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*, Princeton, NJ: Princeton University Press.

[4] Fan, J., and I. Gijbels (1996). *Local Polynomial Modelling and its Applications*, London: Chapman & Hall.

[5] Fan, J., & Yao, Q. (2003). *Nonlinear time series: nonparametric and parametric methods*. New York, NY: Springer New York.

[6] Farmer, L. E., Nakamura, E., & Steinsson, J. (2024). "Learning about the Long Run". *Journal of Political Economy*, 132(10), 3334-3377.

[7] Gáti, L. (2023). "Monetary Policy & Anchored Expectations—An Endogenous Gain Learning Model". *Journal of Monetary Economics*, 140, S37-S47.

[8] Marcet, A., & Nicolini, J. P. (2003). "Recurrent Hyperinflations and Learning". *American Economic Review*, 93(5), 1476-1498.

[9] Milani, F. (2007). "Expectations, Learning and Macroeconomic Persistence". *Journal of Monetary Economics*, 54(7), 2065-2082.

[10] Milani, F. (2011). "Expectation Shocks and Learning as Drivers of the Business Cycle". *The Economic Journal*, 121(552), 379-401.

[9] Marcet and Nicolini (2003) and Milani (2014) are earlier examples of endogenous gains.

[11] Milani, F. (2014). "Learning and Time-Varying Macroeconomic Volatility". *Journal of Economic Dynamics and Control*, 47, 94-114.

[12] Milani, F. (2017). "Sentiment and the US Business Cycle". *Journal of Economic Dynamics and Control*, 82, 289-311.

[13] Patton, A. J., & Timmermann, A. (2010). "Why Do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion". *Journal of Monetary Economics*, 57(7), 803-820.

[14] Slobodyan, S., & Wouters, R. (2012a). "Learning in an estimated medium-scale DSGE model". *Journal of Economic Dynamics and control*, 36(1), 26-46.

[15] Slobodyan, S., & Wouters, R. (2012b). "Learning in a medium-scale DSGE model with expectations based on small forecasting models". *American Economic Journal: Macroeconomics*, 4(2), 65-101.

[16] Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

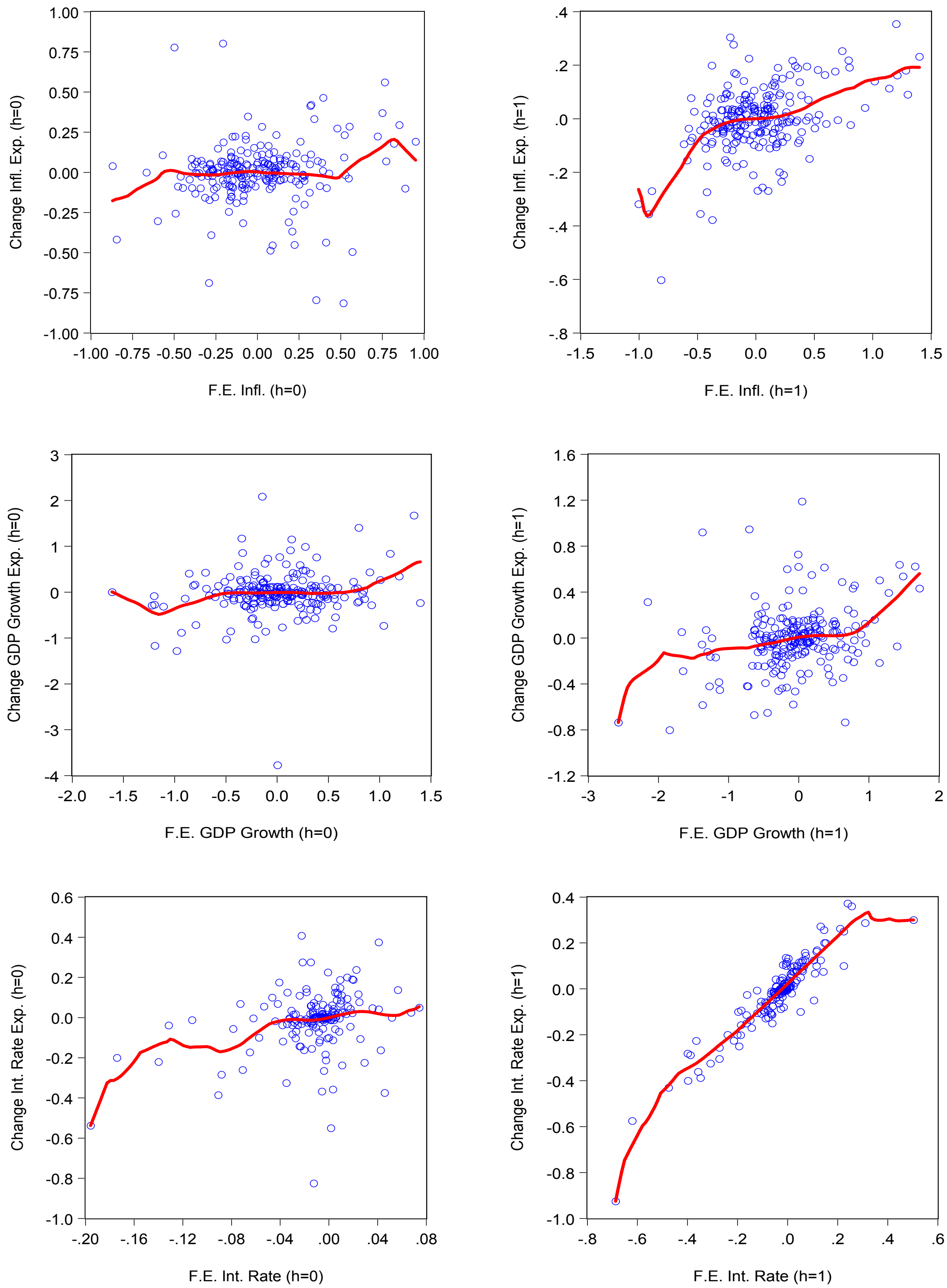

FIGURE 1. Kernel Regressions: Changes in Expectations and Forecast Errors. The top panels refer to inflation expectations ($h = 0, 1$), the medium to output expectations, and the bottom to interest-rate expectations.

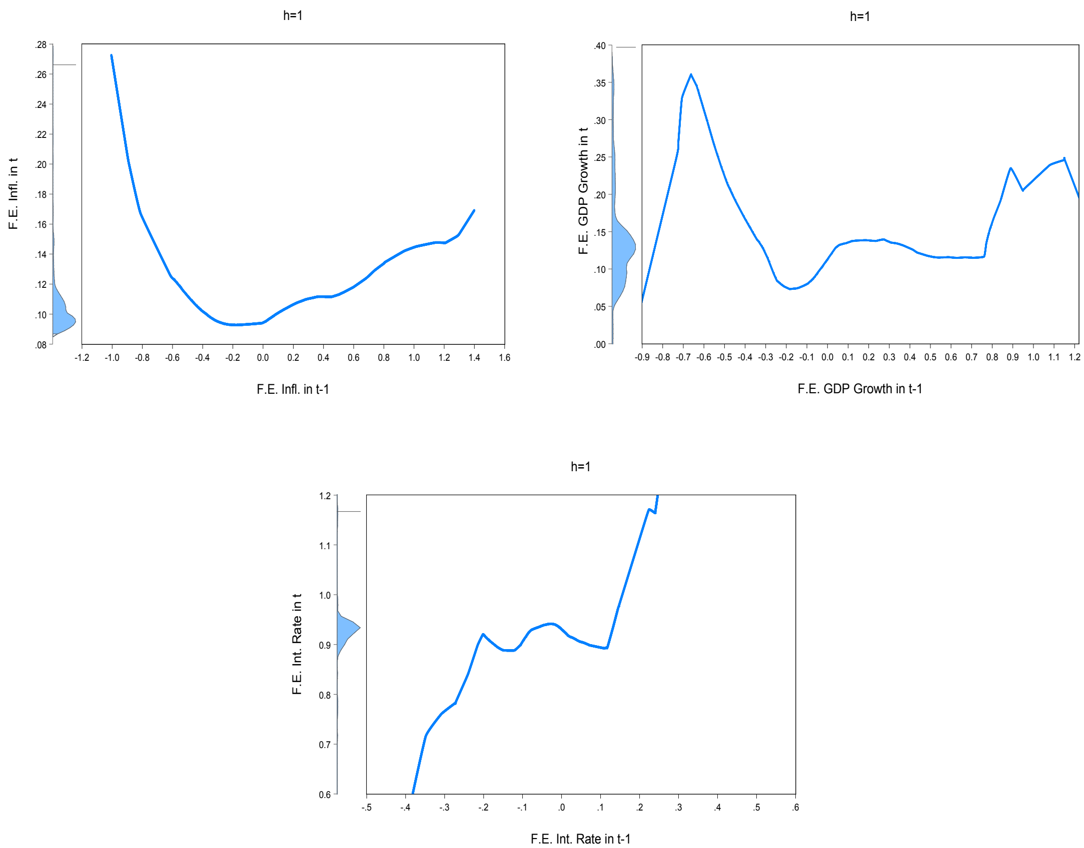

FIGURE 2. Functional-Coefficient Regression Estimates: Changes in Expectations and Forecast Errors. The first panel refers to inflation, the second to output growth, the third to interest rates.