Information, learning, and expectations in an experimental model economy

Michael W.M. Roos*

Wolfgang J. Luhan

Ruhr-Universität Bochum

Lehrstuhl für Makroökonomik and RUBex

D-44801 Bochum, Germany

*(corresponding Author):

Michael.Roos@rub.de

0049 234 3222886

July 10, 2012

Abstract

The experimental "learning-to-forecast" literature finds that subjects use simple linear backward-looking models when forecasting in environments with little to no information about the economic framework. We study the formation of expectations in a laboratory economy of monopolistic firms and labor unions with almost complete knowledge of the model. We observe simple backward-looking rules, but also a considerable share of model-based expectations using information on the economic structure. At least for some subjects, expectations are informed by theory. As in the previous literature, we find individual prediction rules to be heterogeneous.

JEL classification: D83, D84, C99, E00, E30

Keywords: expectations, heuristics, anchoring, model uncertainty, beliefs, mental models, macroeconomic experiment

INTRODUCTION

In this paper we study the formation of expectations in a simple macroeconomic setting using an economic laboratory experiment. Experimental methods have been used before to analyze the formation of expectations (Adam, 2007; Assenza et al., 2011; Bernasconi and Kirchkamp, 2000; Bernasconi et al., 2004; Heemeijer et al., 2009; Hey, 1994; Hommes et al., 2005a,b, 2007, 2008; Pfajfar and Zakelj, 2010; Sonnemans et al., 2004, 2005). Common questions in these papers are whether expectations are rational and how the formation process can be described by simple models. The general result appears to be that expectations are not perfectly rational. Hommes (2011) summarizes the "learning-to-forecast experiments" (henceforth *LTF*) showing that individuals use simple, linear backward-looking models in a number of economic settings and that they switch between different forecasting heuristics (see also Anufriev and Hommes, 2012).

While being very interesting and informative, the conclusions from the LTF experiments may not be generalizable to all kinds of expectations. The particular feature of these experiments is that they consider individual markets with a strong feedback relationship between prices and price expectations. The feedback can be positive as in speculative asset markets or negative as in commodity markets of the cobweb type.

Not all markets have this direct feedback from an expected variable to the realization of the same variable. In differentiated product markets, the demand for an individual producer's products is fairly exogenous for the producer. While there may be negative feedback from the producer's expectation to the price, this situation is better described by learning the demand curve than by learning to forecast, because there is no complex dynamic interaction between the forecasts of several producers and the observed prices that depend on every producer's forecast. This means that coordination, which is a crucial issue in the LTF experiments[1], is not so important in differentiated product markets. Another example are labor markets with wage setters. A union that has the power to set wages should form expectations about the general price level and about the employment response of firms. Both the realized level of employment and the aggregate price level depend only indirectly on the union's expectations. Especially the link between the price level expected by a single union and the realization involves a rather complicated causal chain.

The crucial difference between the examples mentioned above and the markets in the LTF experiments is the degree of stability of the economic environment in which the agents act. In the LTF experiments each agent's forecast has a strong impact on the market which creates complicated dynamics. Under these circumstances it is not surprising that agents switch between different forecasting rules depending on the current state of the market. In other markets the effects of individual actions on the economic environment are weaker and thus the environment subjectively appears to be more stable. Successful forecasting hence does not depend so much on the choice of the best forecasting rule for the current state of the market, but rather may result from acquiring a good understanding of the structure of the economy and using it to form expectations. Under the conditions we have described, it is possible that simple backward-looking expectation formation rules are less important and that expectations are more informed by theoretical perceptions about the model economy.

Our main objective in this study is to learn more about the way subjects form expectations in a macroeconomic context which is characterized by the interaction of different types of agents, weak direct impact of individual expectations on aggregate variables, and feedback from the aggregate to the individual level. In contrast to the LTF literature we are not so much interested in whether there is convergence of individual forecasts to the same (fundamental) value and the properties of the resulting dynamic feedback system. We want to study how subjects perceive the economic environment and how they utilize different kinds of information when forming expectations. For these reasons, our experimental design differs considerably from most of the cited studies, in which subjects have only limited information on the lab economy. In most cases, subjects do not know how the variables to be predicted are determined. If subjects' knowledge about the economic model is very limited, it is difficult to learn and use the true data generating process to form expectations. In Hey (1994), for example, there is not even an economic model that generates the data but only a simple stochastic process. We provide subjects with an almost complete set of information about the model economy's complex structure and the option to buy information about past outcomes and about the actions of other players. This feature allows us to observe which information subjects use to learn about the model. Another common design feature is the separation of expectations from decision making. An exception is a study by Bao et al. (2011) who study expectation formation and decision making jointly to see whether the results from the LTF experiments are robust, if subjects also have to solve an optimization problem. For our purpose, it is natural not to separate decision making and expectation formation, since we want subjects to form expectations about endogenous variables, but no strong link between expectations and realizations. The active decision making can foster experimentation and learning, which might in turn influence their expectation formation.

There are models that explore whether rational expectations equilibria can be learned by agents that behave like econometricians and employ learning methods (see Evans and Honkapohja, 2001 for an overview). Evidence suggests, however, that experts and economic lay people have very different theories of how the economy works (see Roos, 2007; Rubin, 2003) and that lay people in general have poor statistical competencies (see Tversky and Kahneman, 1973, 1974; Kahneman et al., 1982). In such a case, models with statistical learning may not describe well what boundedly rational agents do on a priori grounds. An alternative to the model estimation approach are simple rules of thumb or heuristics that guide the agents' economic behavior, especially when the economic environment is complex and uncertainty is involved, which is typically the case in macroeconomic decision making.

The idea that agents use heuristics to form macroeconomic expectations is not at all new. The adaptive or extrapolative models of expectation formation that were common in Keynesian economics can be interpreted as heuristics. Those models are common alternatives to rational expectations in the experimental literature. Somewhat typical for this literature is the conclusion in Hey (1994) that "subjects are trying to behave rationally, but frequently in a way that appears adaptively"

(p. 329). Another heuristic is the anchoring-and-adjustment heuristic (Tversky and Kahneman, 1974) which has been applied to model expectations in Bao et al. (2012). This heuristic might be a way to connect agents' inclination to use relatively simple rules of thumb with their desire to behave rationally, because it is more flexible than standard forms of extrapolative or adaptive expectations.

Our study is more than a simple test of the rationality of expectations or of the fit of simple adaptive models. While we perform such tests, we want to generate new insights into the process of expectation formation in an explorative manner. Our rich framework allows us to study how subjects deal with complex environments and how they use available information. We find that subjects rarely gather additional information about the economy and form biased expectations. Subjects are heterogeneous in how they form expectations. While simple backward-looking rules are frequent, many expectations can also be explained by other variables from the information sets. Aggregate expectations are not well explained by any single simple model.

Of course, the choice behavior of subjects is as interesting as the formation of expectations, but this is analyzed in a companion paper (Roos and Luhan, 2008). In that paper we find that subjects' perceptions of the model relationships are fairly accurate, but perceptions of other subjects' behavior are biased. Subjects do not optimize for given expectations.

I. MODEL ECONOMY

Our model economy has a deliberately simple structure in order to facilitate the comprehension of the model and its dynamics. Despite the simple structure, the determination of optimal behavior and the formation of expectations are far from trivial.

The economy consists of a labor market and a product market. The labor market determines the nominal wage and the level of employment and the product market determines the prices of goods. Production is directly linked to employment via the production function. The economy consists of three identical industries, each producing a good for which there is a deterministic demand function. Each industry in turn consists of two agents: a monopoly union and a monopoly firm. In a sequential two stage game, the union sets the nominal wage first and the firm subsequently chooses the level of employment for the given wage. In the following, we describe the model in detail. The next section presents the implementation in the laboratory.

In the first stage, in each industry i, a representative worker or union sets the nominal wage w_i. The union's objective function is given by

$$U_i = \delta \left(\frac{w_i}{P} - \frac{\gamma}{L_i} \right), \tag{1}$$

which is a function of the real wage, $\frac{w_i}{P}$, and employment L_i. δ is a scaling factor and γ is the weight given to employment. Since U_i is increasing in the level of employment, it can be interpreted as a utility function of a union that cares for the employment status of its members. The marginal utility of employment is positive, but decreasing. The additive separability of the arguments ensures that the utility function has an inner maximum for the given production function and the product demand function[2].

Next, the firm learns the wage level in the respective industry and determines the industry employment. The profits of firm i are given by

$$\pi_i = p_i Y_i - w_i L_i. \tag{2}$$

Output Y_i is a function of productivity A and the chosen employment level L_i

$$Y_i = A\sqrt{L_i}. \tag{3}$$

The product price p_i is determined by an exogenous and deterministic demand function

$$p_i = \alpha Y_i^{-\frac{1}{\varepsilon}}, \quad \varepsilon > 1. \tag{4}$$

The aggregate price level P is the geometric mean of all industry prices

$$P = \sqrt[n]{\prod_{i=1}^{n} p_i}. \tag{5}$$

Since P depends on the actions of all workers and firms, it is an aggregate variable.

The economy is fully characterized by the equations (1) - (5). As a theoretical benchmark, we derive the equilibrium nominal wage and the equilibrium employment level for homogeneous firms and workers under the assumptions of full information and strict maximization of profits and utility.

The first order condition for fully informed profit maximizing monopolistic firms requires to equate the real producer wage with the markup times the marginal productivity of labor

$$\frac{w_i}{p_i} = \frac{A}{2\sqrt{L_i}}\left(1 - \frac{1}{\varepsilon}\right). \tag{6}$$

Using the production function and the demand function, we derive labor demand as a function of the nominal wage

$$L_i^* = A^{\frac{2(\varepsilon-1)}{\varepsilon+1}} \left(\frac{\alpha(\varepsilon-1)}{2\varepsilon} \right)^{\frac{2\varepsilon}{\varepsilon+1}} w_i^{-2\left(\frac{\varepsilon}{\varepsilon+1}\right)}. \tag{7}$$

Assuming identical industries, the price level in equilibrium is equal to each industry price

$$P^* = \left[\frac{2w\varepsilon\alpha^{\varepsilon}}{A^2(\varepsilon-1)} \right]^{\frac{1}{\varepsilon+1}}. \tag{8}$$

Substituting (7) and (8) into (1) and taking the derivative with respect to w, we obtain the utility maximizing wage

$$w^* = \left(\frac{\alpha^{\frac{\varepsilon}{\varepsilon+1}} A^{\frac{2\varepsilon}{\varepsilon+1}} \left(\frac{\varepsilon-1}{2\varepsilon}\right)^{\frac{2\varepsilon+1}{\varepsilon+1}}}{2\gamma} \right)^{\frac{\varepsilon+1}{\varepsilon}}, \tag{9}$$

which is, of course, constant. In the experiment, we choose the following parametrization: $A = 8$, $\alpha = 1$, $\delta = 5$, $\gamma = 4$, and $\varepsilon = 4$.

II. EXPERIMENTAL DESIGN AND PROCEDURE

In order to introduce an element of uncertainty and to require subjects to form beliefs about the model and expectations about the consequences of their actions, we gave subjects only the rudimentary information of a negative relationship between the output and the price in each industry. We provided neither information about the parameters nor the functional form of the demand function (4).

Equations (1), (2), (3), and (5) were equal for all industries and known to subjects. It was also common knowledge that everybody had this information. In addition to the uncertainty about the model, workers also faced uncertainty about the behavior of firms and consequently about L and p, when setting the nominal wages. While the relationship between p and L is relatively easy to learn, the relation between w and P is complex and intransparent. It appears plausible that real unions only have a limited understanding how their wage setting behavior affects the general price level.

The experiment was implemented computerized using z-tree (Fischbacher, 2007) with networked PCs separated by blinds. Participants were students from different departments of the University of Innsbruck and the University of Oldenburg. Upon arrival in the lab, subjects were randomly assigned to their role as worker or firm and to the economy with fixed assignments throughout the whole experiment (partner design). Instructions (see Appendix) were read aloud and participants were given the opportunity to ask questions before the start of the experiment. During the training phase, participants could check with a test program whether they had understood how their payoffs would be calculated. Only after making sure that all participants had understood how the experiment worked the session was started.

A session consisted of 30 three-stage periods. In the first stage, workers had to choose a nominal wage for the current period from the interval [0, 3]. Simultaneously to the wage decision of the workers, each firm had to enter a wage expectation for the current period[3].

In the second stage firms had to decide, how much labor to employ (from the interval [0.5, 16]) and to state an expectation for their good's price. While firms were making their decision on employment, workers were requested to enter the expectation for the price level P and for the level of employment in their industry[4]. The second stage also lasted 90 seconds.

In the final stage, all subjects learned the realizations of w, L, Y, and p in their industry in the current period. They were also informed about

P, the value of their objective function π or U, their own expectations, and the payoff for their expectations. All subjects were asked if they would like to buy information to be displayed in the first two stages of the next period. If they chose not to buy additional information, only the previous period's realizations of w, L, Y, and p in their industry, P, and the value of their own objective function were displayed. From the second period onward, this information was automatically displayed on the top of the first two stages' screens. In addition to this information, they could opt to buy further information about the market variables (but not about the expectations) in the form of time-series information, cross-section information, or both. In case of the time-series option, they obtained not only the realizations of the variables in the industry of the previous period, but of all previous periods. If they chose the cross-section data, they were shown the previous period's realizations of all variables in all industries. The price of either the time series or the cross-section was 0.1 "taler" (the experimental currency unit, in which both profits and utility was measured) and served mainly as a threshold to deter subjects from constantly requesting all available information.

Two elements contributed to subjects' payoffs in euros. The major part was calculated based on total profits or utility cumulated (both in taler) over all periods. In addition, they received earnings from precise expectations (also in taler). Earnings from expectations were determined by

$$S^f_{ti} = \max(1 - .5(p_t - p^e_{ti})^2 - .5(w_{ti} - w^e_{ti})^2, 0) \tag{10}$$

for firms and

$$S^w_{ti} = \max(1 - .5(P_t - P^e_{ti})^2 - .5(L_{ti} - L^e_{ti})^2, 0) \tag{11}$$

for workers in each period. Forecast earnings were also cumulated over all periods and added to total profit (utility).

We conducted 4 sessions which lasted roughly 2.5 hours each. Each session consisted of 18 participants constituting three economies with three industries in each economy. The average remuneration was €30 per participant including a €5 show-up payment. In each session, we paid out the fixed sum of €540 to be divided conditional on experimental performance. After subtracting €90 paid for the show-up, the remaining €450 were divided by the total number of taler earned to determine the conversion of taler to euros. Subjects received their payoffs in private and in cash directly after the experiment. The payoff scheme was common knowledge and explained in detail before the experiment.

III. RESEARCH QUESTIONS

The objective of this study is to learn how subjects form expectations if they have a lot of structural information about the model economy. In this section we narrow down this fairly general research question to more specific questions for which we will provide specific answers.

The starting point of our analysis is reasoning along the following lines. If subjects had complete knowledge of the model, they would form rational expectations. Since there are no stochastic shocks in the model, the Rational Expectations Equilibrium would be given by equations (7) and (9).[5] In our experiment, however, subjects do not have complete information as they do not know the demand curve.

An economist's approach, which is usually assumed in economic theory, would be to observe the available data and to use it to learn the unknown relationship in order to understand how the model economy works. This is relatively easy here. Although the demand function is not known at the beginning of the experiment, it can be observed both over time and across industries. Subjects knew that all industries were identical. After the first period, they had the option to buy information about the realized values of the variables in the other two industries, including the realized price, output, and employment. This means that in each period after the first one there are at least three price-output combinations available that provide an impression of the demand function's shape. Using the time-series information, subjects can get even more information after a while. Since the demand function does not change over time, the history of the experiment reveals a lot of the missing information. In addition to improving expectations, more information should help making good decisions.

From a bounded rationality perspective, one would not expect that subjects acquire all available information in order to learn the structure of the model systematically. They might not know how to use this information or might not have the cognitive resources to perform the necessary analysis to benefit from the additional data. Our first research question refers to the acquisition of available information:

Question 1: *Do subjects acquire all available information?*

If subjects understand the model structure better over time, we would expect that their expectations become more accurate during the experiment. This could be interpreted as evidence for some kind of learning.

Question 2: *Do expectations become more accurate during the course of the experiment?*

After examining the use and impact of information, we will analyze the properties of the elicited expectations more in detail. One implication of rational expectations is that there is no systematic bias which means that the expected value of the expectation error is zero. The standard test for unbiasedness is to regress the realization of a variable x_t on a constant and the expectation x_t^e and to test for the constant to be zero and the slope to be one (see Lovell, 1986 for a discussion). This is the typical test for Muth's (1961) model of rational expectations. Alternatively, one can test for unbiasedness by running the opposite regression

$$x_t^e = \rho_0 + \rho_1 x_t + \epsilon_t. \tag{12}$$

This model of expectations was called *implicit expectations* by Mills (1957). If the expectation is unbiased then the null hypothesis $\widehat{\rho}_0 = 0$ and $\widehat{\rho}_1 = 1$ cannot be rejected by an F-test.

Question 3: *Are expectations unbiased on average?*

As mentioned in the introduction, expectations can often be described well by simple linear backward-looking models. Standard descriptive models of expectation formation are static expectations

$$x_t^e = x_{t-1}, \tag{13}$$

adaptive expectations

$$x_t^e = x_{t-1}^e + \lambda(x_{t-1} - x_{t-1}^e), \tag{14}$$

and trend extrapolation

$$x_t^e = x_{t-1} + \mu(x_{t-1} - x_{t-2}). \tag{15}$$

We test for these models by estimating the general model

$$x_t^e = \phi_0 + \phi_1 x_{t-1}^e + \phi_2 x_{t-1} + \phi_3 x_{t-2} + \epsilon_t \tag{16}$$

and testing the appropriate restrictions. The restrictions are $\phi_0 = \phi_1 = \phi_3 = 0, \phi_2 = 1$ for the static model, $\phi_0 = \phi_3 = 0, \phi_1 + \phi_2 = 1$ for the adaptive model, and $\phi_0 = \phi_1 = 0, \phi_2 - \phi_3 = 1$ for the trend-following model.

Question 4: *Which simple backward-looking models fit the data best?*

The main question of our paper is whether expectations are also formed by simple backward-looking rules if subjects have a lot of structural information about the economy. It is hard to imagine that subjects who have some knowledge of their economic environment do not use this knowledge at all when they form their expectations. While it may be too extreme to assume that agents use all information (including information about the model) perfectly and form rational expectations, it is likewise extreme to assume the opposite that they do not use any information about relationships between variables. We analyze this hypothesis of *theory-driven expectation formation* by estimating linear models of the form

$$x_t^e = \theta_0 + \theta_1 x_{t-1}^e + \theta_2 x_{t-1} + \Theta\Omega_t + \epsilon_t, \tag{17}$$

where Ω_t is a matrix describing the information set of the subject when the expectation is formed and Θ is a coefficient matrix. We restrict the information set of subjects to those variables that were displayed on each screen. When a subject had not bought additional information, only w_{t-1}, L_{t-1}, p_{t-1}, P_{t-1}, and π_{t-1} or U_{t-1} were displayed. We do not include π_{t-1} and U_{t-1} in the regression model because it is hard to think of a theoretical reason why expectations should depend on the past values of the objective functions[6]. Due to high collinearity between p_t and P_t we only include one of these variables in each estimation. Except for the price expectation model of firms, this is P_t as the price level enters the objective function of the unions. Furthermore, subjects also knew their choice of their decision variables w_t and L_t respectively, when they entered their expectation[7]. We regress the expected values on their on previous value, w_{t-1}, L_{t-1}, p_{t-1} or P_{t-1} and w_t and L_t when they are already known by subjects. We also include the past expectation of the respective variable and the past observation in order to test for the simple backward-looking alternatives of static and adaptive expectations.

Such a general model allows us to test for a purely backward-looking model against some kind of theoretically informed alternative. Even if subjects do not use all information and strictly speaking do not form rational expectations, their expectations can still be driven by theory in the sense that they use some model information to infer relationships between variables. Observing the realizations of those variables they apply the perceived relationships to predict the variables of interest.

Instead of analyzing the objective relationships between the variables in the model economy in a systematic way, subjects might use some kind of heuristic. Hey (1994) found, that many subjects used the previous realization as a starting point and adjusted this value by fractions of previously observed changes. This finding suggests that subjects use a variant of the anchoring-and-adjustment heuristic. The adjustment process might work as follows. We assume subjects to have mental models of the experimental environment that specify perceived causal relationships between variables. These mental models are partly derived from the experimental instructions, which contain all model equations except for the demand function. But the information from the instructions is not sufficient to obtain the necessary causal relationships, even if it were complete. The instructions contain only structural equations, not the reduced forms. Furthermore, the model is only helpful with the auxiliary assumption of rational players. For this reason, subjects modify and extend the model as it is presented in the instructions to mental models that appear plausible and useful. Mental models are useful if they allow subjects to make satisfactory choices and predictions that are not utterly wrong. Of course, we cannot observe those mental models, but it might be possible to learn something about those models from the variables that help to explain expectations.

Question 5: *Are expectations influenced by other variables and theoretical considerations as opposed to simple backward-looking models?*

IV. RESULTS

Information

Our first research question focuses on information acquisition. Given a fixed yet unknown demand function (4) the available information would enable participants to learn the shape of the function thereby improving expectations as well as decisions.

We find a rather low demand for information: Only in 26.2% of all cases any information was purchased; in other words, almost 3/4 of the stated expectations (and the decisions taken subseqently) were formed without the use of the additional information available. We find a strong preference for the cross section information. In 83.7% of all information purchases subjects requested either the cross section alone (58.5%) or in combination with the time series information (25.2%).

Observing this low demand for information the question arises of whether there is a common pattern in information purchases valid for all subjects. One reason for the low information requests could be that subjects consider the information not useful and therefore disregard it when forming expectations and choosing w and L respectively. Should their experiences from early information purchases lead them to believe that the information cost of 0.1 taler exceed the returns, we would observe a decrease in information purchases over time.

Instead of a common pattern for all subjects we identify 18 out of the 72 subjects whose information requests account for 67.8% of all information purchases (Kolmogorov-Smirnov-Test $p < 0.001$ comparing the frequency of information purchases). We classify subjects as "information buyers" if they requested information in the majority (> 15) of periods. Within this group of information buyers we do not observe a decrease of information purchases over time but rather a stable pattern throughout the periods. The remaining subjects buy information only infrequently without any observable pattern or trend. It appears that those subjects who bought the information valued them higher than the costs. But do we actually find increased earnings due to information purchases, i.e. is the value of information higher than the cost? A first glance at the descriptive statistics confirms the value of information: Both unions (8 subjects) and firms (10 subjects) in the information buying groups earned above average payoffs. While the overall average period payoff was 4.3 taler for unions and 2.6 taler for firms, the unions that frequently bought information earned on average 5.6 taler and the firms of that group earned on average 3.1 taler per period. Although the payoffs are significantly higher for both types of information users ($p = 0.019$ for workers and $p = 0.029$ for firms in a two-tailed t-test), we find a significant increase in the earnings for predictions, S, only for the participants in the role of firms ($p=0.003$ in a two-tailed t-test), but not for workers ($p = 0.128$). This result is rather puzzling as improved predictions of P, L, and p would increase S as well U and π.

To clarify the picture we estimate fixed effect panel models for S_t as reported in the first column of Table 1, using Driscoll and Kraay (1998) standard errors to control for autocorrelation and potential cluster effects[8]. The first regressor is information purchase in the previous period ($info_{t-1}$), as this determined the information available when eliciting expectations and making decisions[9]. *firm* is a dummy variable for firms and accounts for differences in S depending on the subject's role in the experiment. With *period* we model a time trend and $period^2$ allows this trend to be non-linear. A significant trend in time might be interpreted as learning and serves as first observation regarding research question 2.

Table 1 here.

In line with the results above, we do not find a significant impact of the information purchase on the expectation accuracy (S_t). Firms do, ceteris paribus, perform better in predicting expected values (column 1) which might be driving the non-parametric results of a significant impact of information on S_t.

We report a significantly positive, non-linear time trend for S_t which suggests gradual improvement of the expectation formation and can be interpreted as a first indication of learning. For a more elaborate investigation of research question 2 we calculate the **r**elative **a**bsolute **e**rror (rae) for all predicted variables (L, P, w, p) to measure the accuracy of expectations during the course of the experiment. The relative absolute prediction error ($x_t^{rae} = \mid x_t^e - x_t \mid / x_t$) is normalized by the realization of the variable to be predicted and therefore comparable across the different variables. Figure 1 depicts the period means for all prediction errors with the workers' prediction errors (L^{rae}, P^{rae}) in the upper and the firms' prediction errors (w^{rae}, p^{rae}) in the lower part of the figure. A first visual examination reveals diminishing prediction errors over time for all four variables, indicating that subjects learn to make better predictions over time. Two observations are quite eye-catching: First, the firms perform better in predicting both of the variables. Second, the prediction of the price (p) as well as the price level (P) appears to be easier than predicting the wage or labor demand. The explanation might be that the prices result from an unknown but deterministic demand function while wages and labor demand are set by other players. According to Figure 1 the participants found it harder to learn their counterparts' strategies than to learn a part of the model.

Figure 1 here

In order to quantify and to test the effects we see in Figure 1, we estimated the relative absolute errors as functions of the information purchase in the previous period and a quadratic time trend. The results can be found in columns 2 to 5 of Table 1. The estimation results confirm a negative, non-linear time trend in all prediction errors but no significant impact of information acquisition.

Summing up, we find that most subjects acquire little to no information when making decisions and forming expectations. If they do, this does not improve the accuracy of their predictions. Participants who frequently buy information do indeed earn higher profits. The evidence suggests, however, that there is no causal relationship between information acquisition and performance. The predictions do become more precise over time which we interpret as evidence of learning.[10] This learning, however, must be some kind of learning from experience as it depends on time, not on information. With respect to decisions, learning occurs through better expectations. Those subjects with more precise expectations earn more due to better decisions.

Unbiasedness

Since we are primarily interested in the aggregate properties of the data we estimated equation (12) by pooled OLS regressions[11]. We consider this procedure as an aggregate of the individual estimations, which are of minor interest in the context of the formation of macroeconomic expectations.

Two econometric problems can arise in this kind of analysis. First, the residuals might be autocorrelated, if there is some kind of adaptive behavior of individuals. Second, there is reason to believe that the observations within each economy are not independent, because there is one common price level in each economy that affects all workers in an industry and because all members of an economy had the possibility to observe what the other members did. While we do not think that there is strong correlation between the observations in each industry, we nevertheless correct for cross-sectional correlation and autocorrelation by using Driscoll and Kraay (1998) robust standard errors.

All aggregate expectations are clearly biased as Table 2 shows. For all variables the F-tests reject unbiasedness of expectations at the 1% level. A basic condition for rational expectations is therefore generally violated in our observations.

Table 2 here

Given that most subjects acquired little additional information about the economy, it is not surprising to find that in general expectations are biased. But not even those subjects who frequently bought additional data were able to avoid bias. We sorted subjects into two groups according to the frequency of their information purchases (information buyers as defined in the previous subsection) and tested for bias in both groups separately. For p, w, and L the estimated parameters are not statistically different between the two groups (Wald test). Only for the price level P the Wald test detects a significant difference at the 2%-level. Subjects who acquired information regularly produced unbiased expectations P^e. Performing the same regressions and tests at the individual level we find that of the 36 subjects in the role of firms, the price expectations and the wage expectations of 8 subjects are unbiased. Workers' expectations of the employment level are unbiased in 13 cases and price level expectations in 6 cases.

We have found evidence that aggregate expectations for all four variables in question are generally not rational due to bias. At the individual level, only a minority of about one quarter of the subjects form unbiased expectations. Whether subjects acquire additional information or not hardly matters for the unbiasedness of expectations.

Backward-looking models

The estimation[12] of model (16) allows us to test whether simple linear backward-looking models fit the pooled data. Table 3 shows that only in the case of employment expectations adaptive expectations as a standard model fits the pooled data. For all other variables none of the simple models fit the data according to the F-tests[13] even though the adaptive model seems to be a fair description of wage expectations with respect to the point estimates of the individual coefficients. In the aggregate there is little evidence for static expectations or trend following behavior. The coefficient on the second lag is only significantly different from zero in column (1), where it is positive and the coefficients imply trend-reverting rather than trend-following behavior.

Table 3 here

As discussed above, a large part of the LTF literature reports heterogeneity concerning the forecasting rules. We therefore estimate model (16) also for each individual and test for the restrictions of the simple expectation models[14]. Our results[15] show that there is heterogeneity both across individuals and across variables (see Table 4). For price expectations, two subjects form static expectations, 4 subjects form adaptive expectations and one subject uses a trend rule. For the 20 subjects that are classified as "other" at least one of the past variables matters, but not in a way consistent with one of the standard models. The price expectations of 9 subjects cannot by explained at all by p^e_{t-1}, p_{t-1} and p_{t-1}. The most frequent model of individual wage expectations is the adaptive model, followed by the static model (14 and 6 subjects). Again, a trend rule fits only the expectations of one subject. Qualitatively, the picture is similiar for the employment and price level expectations of the workers. The most frequent model is the adaptive one, followed by the static model. But again, for a large number of subjects expectations cannot be explained by the considered backward-looking models.

Table 4 here

We conclude that adaptive expectations is the best simple backward-looking model, both in the aggregate and at the individual level. However, there is considerable heterogeneity and many individual expectations cannot be described well by one of the simple models.

Theory-driven expectations

Finally, we estimate models augmented by other variables included in subjects' informations sets to allow for more model driven forecasting behavior. Table 5 contains the results of the estimation of (17) for all subjects pooled together. The most important result is that - if we allow for more informed, model-driven learning - a purely adaptive model of the type $x_t^e = \theta_1 x_{t-1} + \theta_2 x_{t-1}^e + \Theta\Omega_\mathbf{t}$ with $\theta_1 + \theta_2 = 1$ and $\Theta = 0$ is strongly rejected as shown by the F-tests in all cases. Expectation formation in our experiment cannot be described well by a mechanistic adaptive model, which only uses past information of the variable and its expectation. For w^e the restriction $\theta_1 + \theta_2 = 1$ is rejected, but no other coefficients are different from zero (see column (2)). In all other regressions, at least one element in Ω_t is significantly different from zero which implies that expectations in our experiment are not purely backward looking.

Table 5 here

We again estimate all models at the individual level and classify subjects into different groups (see Table 6). If x_{t-1} is not different from one and all other coefficients are zero in individual t-tests, the best model is static expectations. Static expectations do not describe well price expectations, but between 6 and 9 subjects use this model to predict the other variables.

If the sum of coefficients of x_{t-1} and x^e_{t-1} is one according to an F-test and all other variables are zero in t-tests, the model is adaptive. Adaptive expectations describe the wage, employment, and price level expectations of 4, 5, and 2 subjects respectively. Notice that these numbers are considerably lower than in the previous section where the estimation model did not include other variables.

If x_{t-1} is not different from one and at least one other variable is significantly different from zero, we call the model anchoring-and-adjustment model. This model is the most frequent model for price expectations (10 subjects), employment expectations (7 subjects), and price level expectations (7 subjects). Wage expectations are formed by 3 subjects in this way. An illustrative example of such an anchoring-and-adjustment rule is

$$L^e_t = \underset{(.30)}{0.15} + \underset{(.02)}{0.98} L_{t-1} - \underset{(.15)}{0.7} w_t + \underset{(.22)}{0.75} w_{t-1}$$

which describes well the expectations of three subjects ($R^2 = 0.98$, 81 observations)[16]. These three subjects assume that the level of employment will not change unless they change the nominal wage.

In addition to these heuristic prediction rules, we find that subjects often use rules that seem to be driven by theoretical considerations or at least can be rationalized easily. Subjects knew that there was a negative relation between the produced quantity and the price of a good. We hence classify all individual price expectation models as theory driven if we found a significantly negative coefficient on employment which was the case for 8 subjects[17]. A theory-driven expectation might be that firms predict a positive relation between w_t and P_{t-1} because they might believe that workers want to make up for high price levels in the past by higher nominal wages. This is a common justification for nominal wage increases in real-world wage negotiations. In fact, the wage expectations of 3 subjects exhibit this feature. Workers sensibly expect that there is a negative relation between the wage they set and the employment level chosen by the firm. If the coefficient on w_t is significantly negative in the employment expectation regression, we classifiy this model as theory driven. Since we do not see a simple and convincing story for theory-driven expectations of P^e we do not sort any price level expectation models in this group. As before a large number of individual level expectations do not fall into one of the mentioned categories or cannot be explained by the estimated model at all.

Table 6 here

We find that in many cases subjects seem to have used more sophisticated models than the simple backward-looking rules common in the literature. The expectations of many subjects can be characterized by an anchoring-and-adjustment rule that uses the last observation as the anchor and some other relationship as the adjustment rule. There is, however, a large degree of heterogeneity with respect to the method used for the adjustment and with respect to the expectation formation rules in general. We do not find any dominant models at the individual level which implies that the models estimated at the aggregate level are mixtures of different individual models.

V. CONCLUSIONS

Previous experimental research on the formation of expectations found that expectations are typically not rational and often are well described by simple backward-looking rules such as static or adaptive expectations. Another established result in the experimental literature is that there is a lot of heterogeneity in how individuals form expectations.

Our experimental design differs from those other researchers have used in that the subjects in our experiment had to make decisions and form expectations in a model with monopolistic labor and product markets. There is uncertainty in our model, not because of exogenous stochastic shocks, but because subjects are not perfectly informed about the model structure. Nevertheless, they have plenty of structural information available and can learn the unknown relationship over time quite precisely by using observations from the experimental history. We use this particular design in order to analyze whether in an information-rich environment expectations are still best described by simple backward-looking rules.

We find that subjects' expectations are clearly not rational, because the fundamental requirement of unbiasedness is systematically violated. We also observe that subjects acquire little additional information that would have helped them to learn the structure of the model. Furthermore, those subjects who did gather the information do not benefit from it. Previous information acquisition has no significant impact on the accuracy of expectations. Nevertheless, subjects' expectations become better over time, which might be an indication of learning. This learning, however, seems to be driven rather by direct experience than by active exploration of the model. Our findings are hence more favorable to heuristic behavior than to systematic learning.

Confirming the results in the LTF literature, our estimated expectation models differ widely at the individual level which Hommes (2011) calls the heterogeneous expectations hypothesis. Among these individual models, adaptive expectations and static expectations are quite frequent. However, the main contribution of our paper is to show that the expectations of a significant number of subjects are better described by more sophisticated models taking structural information about the economy into account. One kind of these expectation models is what we call the anchoring-and-adjustment model that takes the last observation of the variable to be forecast as an anchor and adjusts from there using additional information. For all variables except for the wage, the anchoring-and-adjustment model is among the most frequently used models. But some subjects' expectations are even more informed by theory. Price expectations, employment expectations and to a smaller extent also wage expectations are formed in a manner that is consistent with theoretical considerations. About one fifth of the subjects' price expectations depend negatively on the level of employment and employment expectations depend negatively on the wage level which is also predicted by the theoretical model.

Our results fit well into the concept of bounded rationality in the sense that subjects behave in a sensible way, but fail to achieve full rationality. Many of our subjects do not resort to mechanistic forecasting rules that can be applied in any situation, but use structural information about their environment. This means that they attempt to understand how the variables in the model economy are determined. However, as the large variety of different forecast models at the individual level shows, subjects seem to have heterogeneous mental models which may be very different from the actural model underlying the experiment. Their behavior is also boundedly rational in the sense that most subjects do not explore the model economy in a systematic way which we conclude from the fact that they acquire little information in addition to the information they already have.

With heterogeneous expectations that are based on different subjective models in a relatively simple laboratory experiment it appears very unlikely that macroeconomic expectations outside the laboratory can be modelled adequately by a single simple model. We conjecture that aggregate expectations about macroeconomic variables are the result of a diverse mix of models, from simple mechanistic ones to rather sophisticated variants informed by some kind of theory. It is a challenge for future research to explore the properties of such aggregate expectations which are derived from heterogeneous individual expectations.

AKNOWLEDGEMENTS

We would like thank the editor and two anonymous referees for their excellent comments. We would also like to thank conference participants at the 1^{st} LeeX International Conference on Theoretical and Experimental Macroeconomics (Barcelona, 2010) and the 3^{rd} Nordic Conference on Behavioral and Experimental Economics (Copenhagen, 2008) as well as seminar participants at the University of Aachen (2007) for their comments. Wolfgang Luhan gratefully aknowledges financial support from the University of Oldenburg.

Notes

[1]In LTF experiments the price is usually a function of the average price expectation, so that stable price dynamics require coordination on the same price expectation mechanism. Whether this coordination occurs is one of the questions in the LTF literature (e.g., Hommes et al. 2005b, Heemeijer et al. 2009).

[2]This function may not be the standard utility function in union models, but it confronts workers with the intended trade-off between the real wage and the employment level. We choose this particular functional form simply for analytical convenience.

[3]We explain below, how firms' expectations were incentivized, see equation 10.

[4]See equation 11 below.

[5]Our parametrization yields $w^*_{fi} = 0.523$ and $L^*_{fi} = 2.524 \cdot {w^*_{fi}}^{-1.6} = 0.640$.

[6]In control estimations we did not find significant effects of these variables.

[7]With the exception of w_t^e which had to be entered by firms before they made their employment choice.

[8]It is obvious from (10) and (11) that S_t is censored at the lower limit of 0, which would suggest a panel-tobit-estimation. As this does not change the result, we report the panel-OLS for both dependent variables.

[9] $info_{t-1}$ is a dummy variable that marks if any information has been bought independent of the type. We tested alternative models including the type and quantity of information bought as well as the cumulated information bought up to period t. These alternative regressors were neither significant nor did they improve the model fit.

[10]Decreasing forecasting errors could in principle be also due to an exogeneous stabilization of the model dynamics. We thank an anonymous referee for pointing out this alternative interpretation.

[11]We dropped the first three observations in order to allow for some initial learning. We also exclude observations where $p^e > 1$ or $P^e > 2$, as they are obvious outliers. Actual prices were never greater than 1 and most of these cases are clearly identified as mistakes, e.g. typos like 55 instead of 0.55.

[12]We estimate by OLS and correct the standard errors for cluster effects. In the estimation we eliminate the observations of the first three periods and some obvious outliers for the price and the price level. See footnote 10.

[13]From equation (16) we derived the following restrictions: static model $\phi_0 = \phi_1 = \phi_3 = 0, \phi_2 = 1$; adaptive model $\phi_0 = \phi_3 = 0, \phi_1 + \phi_2 = 1$; trend-following model $\phi_0 = \phi_1 = 0, \phi_2 - \phi_3 = 1$.

[14]In several cases the F-tests did not reject several of the three sets of restrictions. In this case we also used individual t-tests to decide which one of the three models

is most appropriate by assuming that a cofficient is zero if this was not rejected by the t-test.

[15]All results at the individual level not shown here are available upon request from the authors.

[16]Except for the constant, all reported coefficients are significantly different from zero in t-tests at 5%. The values in parentheses are robust standard errors.

[17]Other variables can be different from zero as well as long as it does not interfere with the achoring-and-adjustment rule.

REFERENCES

Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *Economic Journal*, **117**, 603-635.

Anufriev, M. and C. Hommes (2012). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, forthcoming

Assenza, T., P. Heemeijer, C. Hommes, and D. Massaro (2011). Individual expectations and aggregate macro behavior. CeNDEF Working Paper, University of Amsterdam.

Bao, T., J. Duffy, and C. Hommes. (2011). Learning, forecasting and optimizing: An experimental study. CeNDEF Working Paper, University of Amsterdam.

_, C. Hommes, J. Sonnemans, and J. Tuinstra. (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, forthcoming.

Bernasconi, M., and O. Kirchkamp (2000). Why do monetary policies matter? An experimental study of saving and inflation in an overlapping generations model. *Journal of Monetary Economics*, **46**, 315-343.

_,_ and P. Paruolo (2004). Do fiscal variables affect fiscal expectations? Experiments with real world and lab data. *Journal of Economics Behavior and Organization,* **70**, 253-265.

Driscoll, J., A. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, **80**, 549-560.

Evans, G., S. Honkapohja (2001). *Learning and Expectations in Macroeconomics.* Princeton University Press, Princeton.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics,* **10**, 171-178.

Heemeijer, P., C. Hommes, J. Sonnemans, J. Tuinstra (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control* 33: 1052-1072.

Hey, J.D. (1994). Expectations formation: Rational or adaptive or ...? *Journal of Economic Behavior and Organization,* **25**, 329-349.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, **35**, 1-24.

_, J. Sonnemans, J. Tuinstra, and H. van de Velden (2005a). A strategy experiment in dynamic asset pricing. *Journal of Economic Dynamics and Control,* **29**, 823-843.

_, _, _,_ (2005b). Coordination of expectations in asset pricing experiments. *Review of Financial Studies,* **18**, 955-980.

_, _, _,_ (2007). Learning in cobweb experiments. *Macroeconomic Dynamics,* **11**, 8-33.

_, _, _,_ (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior and Organization,* **67**, 116-133.

Kahneman, D., P. Slovic, A. Tversky (1982). Judgment under uncertainty: Heuristics and biases. Cambridge, Cambridge University Press.

Lovell, M.C. (1986). Tests of the rational expectations hypothesis. *American Economic Review,* **76**, 110-124.

Mills, E. (1957). The theory of inventory decisions. *Econometrica,* **25**, 222-238.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica,* **29**, 315-335.

Pfajfar, D., B. Zakelj (2010). Inflation expectations and monetary policy design: Evidence from the laboratory. Working paper, Tilburg University.

Roos, M. W. M. (2007). Nonexpert beliefs about the macroeconomic consequences of economic and noneconomic events. *Public Choice*, **132**, 291-304

_, W. Luhan (2008). As if or what? - Expectations and optimization in a simple macroeconomic environment. Ruhr Economic Paper No. 55.

Rubin, P.H. (2003). Folk economics, *Southern Economic Journal,* **70**, 157-171.

Sonnemans, J., C. Hommes, J. Tuinstra, and H. van de Velden (2004). The instability of a heterogeneous cobweb economy: a strategy experiment on expectation formation. *Journal of Economic Behavior and Organization*, **54**, 453-481.

_, P. Heemeijer, C. Hommes (2005). Price expectations in the laboratory in positive and negative feedback systems. paper presented at Computing in Economics and Finance 2005 165, Society for Computational Economics.

Tversky, A., D. Kahneman (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* **5**, 207-232.

_, _ (1974). Judgment under uncertainty: Heuristics and biases. *Science,* **185**, 1124-1130.

Tables and Figures

Table 1: Information impact and learning, panel regression

	(1)	(2)	(3)	(4)	(5)
	S_t	L^{rae}	P^{rae}	w^{rae}	p^{rae}
cons	-	2.095**	.202**	.279**	.215**
		(.365)	(.166)	(.010)	(.031)
$info_{t-1}$	.004	.159	-.005	.002	-.001
	(.007)	(.089)	(.006)	(.007)	(.005)
firm	.293**				
	(.029)				
period	.043**	-.146**	-.014**	-.014**	-.019**
	(.005)	(.0045)	(.002)	(.002)	(.003)
$period^2$	-.001**	.003**	.0002**	.0003**	.0005**
	(.000)	(.001)	(.000)	(.000)	(.000)
R^2	.126	.033	.046	.025	.181
#obs	2088	1044	1027	10441	944

fixed effects, robust standard errors (Driscoll-Kraay) in parentheses, **, * indicate significance at the 1% and 5% level respectively.

Table 2: Biasedness of expectations, pooled samples

	(1)	(2)	(3)	(4)
x_t^e	p^e	w^e	L^e	P^e
cons	.15**	.58**	1.06*	.23**
	(.04)	(.06)	(.16)	(.07)
x_t	.73**	.64**	.75**	.59**
	(.06)	(.03)	(.03)	(.12)
R^2	.37	.48	.44	.05
$p(F)$	.001	.000	.000	.007
#obs	926	972	972	956

robust standard errors (Driscoll-Kraay) in parentheses, $p(F)$: significance level of F-test on $\widehat{\rho}_0 = 0$ and $\widehat{\rho}_1 = 1$, asterisks indicate significant difference from $\widehat{\rho}_0 = 0$ and $\widehat{\rho}_1 = 1$ in t-test at 5% and 1% for * and ** respectively

Table 3: Backward-looking models, pooled samples

	(1) p^e	(2) w^e	(3) L^e	(4) P^e
x^e_{t-1}	.00	.42**	.38**	.05
	(.00)	(.09)	(.11)	(.03)
x_{t-1}	.54**	.57**	.68**	.65**
	(.05)	(.03)	(.06)	(.07)
x_{t-2}	.19**	-.05	-.11	.07
	(.04)	(.07)	(.09)	(.13)
constant	.14**	.07	.31*	.13
	(.02)	(.05)	(.11)	(.08)
R^2	.345	.828	.706	.125
$p(F)_{static}$	.00	.00	.00	.00
$p(F)_{adaptive}$	.00	.01	.10	.00
$p(F)_{trend}$	.00	.00	.00	.00
# obs	926	972	972	956

robust standard errors in parentheses, $p(F)$: significance level of F-tests, asterisks indicate significant difference from zero in t-test at 5% and 1% for * and ** respectively

Table 4: Backward-looking models, individual results

	p^e	w^e	L^e	P^e
static	2	6	8	6
adaptive	4	14	9	9
trend	1	1	3	2
other	20	14	12	11
not explicable	9	0	4	10

Entries are the numbers of subjects for which the respective models fit.

Table 5: Models of expectations, pooled samples

	(1) p^e	(2) w^e	(3) L^e	(4) P^e
p_{t-1}	.71**			
	(.17)			
p^e_{t-1}	.00			
	(.00)			
w_t	.03**		-.44	.01*
	(.01)		(.34)	(.003)
w_{t-1}	-.03**	.55**	.60**	.01
	(.01)	(.04)	(.20)	(.01)
w^e_{t-1}		.37**		
		(.05)		
L_t	-.01**			
	(.001)			
L_{t-1}	.01	-.002	.68**	-.001
	(.004)	(.005)	(.09)	(.001)
L^e_{t-1}			.32**	
			(.08)	
P_{t-1}		.65	2.95	.54*
		(.38)	(1.47)	(.19)
P^e_{t-1}				.05
				(.03)
$const$	.16	-.25	-1.68	.19*
	(.09)	(.21)	(.83)	(.08)
$\overline{R}^2$	.424	.829	.713	.150
$p(F)_{adaptive}$	.000	.020	.01	.023
$\#obs$	926	972	972	956

robust standard errors in parentheses, $p(F)$: significance level of F-tests, asterisks indicate significant difference from zero in t-test at 5% and 1% for * and ** respectively

Table 6: Theory-driven expectations, individual results

	p^e	w^e	L^e	P^e
static	0	9	6	7
adaptive	0	4	5	2
anchoring	10	3	7	7
theory-driven	8	3	7	
other	13	14	8	10
not explicable	5	3	3	10

Entries are the numbers of subjects for which the respective models fit.

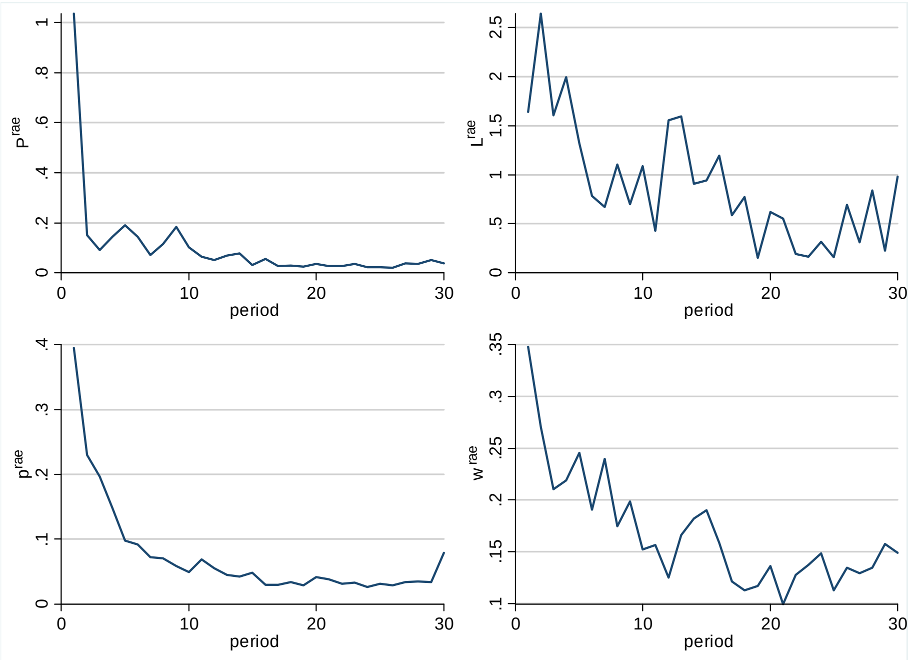

Figure 1: Relative absolute prediction errors, period means