Learning Trend Inflation – Can Signal Extraction Explain Survey Forecasts?*

Steffen Henzel†

November 11, 2008

Abstract

It can be shown that inflation expectations and associated forecast errors are characterized by a high degree of persistence. One reason may be that forecasters cannot directly observe the inflation target pursued by the central bank and, hence, face a complicated forecasting problem. In particular, they have to infer whether the observed movement of the inflation rate is due to a permanent change of policy parameters or whether it is the result of a transient shock. Consequently, it is assumed that agents behave like econometricians who filter noisy information by estimating an unobserved components model. This constitutes the trend learning algorithm employed by the forecaster. To examine whether this is a valid assumption, I fit a simple learning model to US survey expectations. The second part contains an out–of–sample forecasting experiment which shows that learning by signal extraction matches survey measures closer than other standard models. Moreover, it turns out that a weighted average of different expectation formation processes with a prominent role for signal extraction behavior is well suited to explain survey measures of inflation expectations.

JEL classifications: E31; E37; C32; C53
Keywords: Inflation expectations; State–space model; Learning; Signal extraction

*I would like to thank Gebhard Flaig, Kai Carstensen, participants at the European meeting of the Econometric Society 2008 as well as the participants at a research workshop at the University of Munich for valuable suggestions and helpful comments on an earlier version of this paper.

†Ifo Institute for Economic Research at the University of Munich,
Poschingerstr. 5, 81679 Munich, Germany, Email: Henzel@ifo.de

1 Introduction

Perfect foresight and full information about constant policy parameters are common assumptions in many macroeconomic models. However, one implication of the standard New Keynesian model with rational expectations is that an unanticipated change in the inflation target would lead to a sudden jump in the level of inflation expectations. Consequently, if the model is purely forward–looking, disinflation is not accompanied by an output loss. However, both implications, the jump behavior of expectations and the absence of disinflation cost can certainly be doubted. Focussing on inflation expectations here, these can be shown to exhibit a series of features that are inconsistent with the assumption of rational expectation formation in the sense of Muth (1961). For instance, Evans and Wachtel (1992) find that U.S. inflation expectations are biased and inefficient predictors of future inflation.[1] In particular, forecast errors are found to be persistent and can be explained ex–post.

In their paper, Evans and Wachtel (1992) emphasize the empirical relevance of information constraints for the formation of inflation expectations. They state that the forecasts generated by their univariate regime–switching model exhibit some important properties of survey data on inflation expectations. A more recent series of papers also relaxes the assumption of perfect knowledge. Among others, Kozicki and Tinsley (2005), Andolfatto and Gomme (2003), Nunes (2004) and Erceg and Levin (2003) emphasize the importance of the persistence of expectations for the inflation process. One of their findings is that learning behavior is important to explain the transition dynamics of monetary policy which has implications for the design of monetary policy.[2] They also stress that this has resulted in quantitatively important welfare effects on output and interest rates during disinflation episodes. In these papers, inflation persistence stems from the fact that rational agents face a complicated forecasting problem when forming expectations. Theoretically, one reason is that they only observe noisy information, which constitutes a signal extraction problem. Like in Cukierman and Meltzer (1986), the problem arises from the fact that it cannot be distinguished between permanent target shocks and transitory shocks to the policy rate. Consequently, if the conduct of monetary policy changes over time, thereby changing the inflation target, agents face a complicated forecasting problem: the decomposition of inflation into trends and transitory components. The investigation of trend breaks in measures of inflation expectations has rarely been a subject of studies. However, the behavior of expectations following a trend shift is of importance. As noted above, sluggish expectations will constitute

[1] Among others see Roberts (1997) and Branch (2004) and the papers cited there.

[2] For instance, Andolfatto and Gomme (2003) advocate the idea that it is important for a central bank to be credible, as this significantly reduces the output inflation trade–off. For other implications of learning for an optimal monetary policy see Evans and Honkapohja (2003). Also see Cogley and Sbordone (2006) on implications for the New Keynesian Phillips curve.

the persistence of inflation rates. From the point of view of the monetary authority, it is important to know what implications a trend shift has on inflation expectations because it will determine how costly a disinflation policy will be in terms of output loss, as emphasized by Nunes (2004). Andolfatto and Gomme (2003), for example give a theoretical justification why transparency will reduce the cost of disinflation.

Here, I basically follow the above frameworks and assume that agents have to make decisions in an environment that is characterized by noisy information. This, in turn, leads to a forecasting problem for decision makers which is more complicated than in perfect foresight models: How can the action of a central bank be interpreted in the light of new information? Does a shift of the inflation rate stem from a temporary policy action or some other temporary shock or does it reflect a permanent change in the policy parameters – i.e. the inflation target? Certainly, agents have to form expectations about these issues which will ultimately be reflected in their projections of the inflation rate. A lack of information in this context arises for different reasons. Either the monetary authority refrains from complete disclosure of the policy making process (intransparency) or the announced inflation target may not be fully credible. Another theoretical explanation for the rejection of rational expectations is that agents are boundedly rational, which means that they face resource constraints if information is costly (rational inattention) or that they lack sophistication. This leads Branch and Evans (2006) to emphasize the importance of simple forecasting models on the part of private agents to model expectations. Nevertheless, rational agents would, within these frameworks, solve the forecasting problem by application of some (optimal) learning algorithm. To account for bounded rationality, I will assume a *simple* forecasting model on part of private agents. Following the approach of Branch and Evans (2006) and Dossche and Everaert (2005), I will restrict the analysis to univariate forecasting models. The advantage is that no assumptions about the structural relationships between variables have to be made and results will not depend on a specific theoretical model.

In the literature on learning, it has become standard to assume that agents act like econometricians who estimate the unknown parameters of the forecasting model from past data. As outlined above, here, they are also confronted with a signal extraction problem they solve by estimating permanent and transitory components of inflation in order to come up with a forecast. In particular, the forecasters will update trend perceptions each time a new observation becomes available. To do so in an efficient way, they build up an unobserved components model and make use of the Kalman filter recursions.

The remainder of the paper is organized as follows. In section 2 I will have a closer look at the different survey measures for inflation expectations. Thereby I will briefly update and review some of the inherent characteristics of inflation expectations. Section 3 investigates whether a model of signal extraction can be fit to survey measures for U.S. inflation expectations, thereby answering the question: Do agents update trend expectations in the light of past forecast errors? At this point, it will also be of interest to know how long it takes until agents learn about a new monetary policy regime. Taking an out–of–sample perspective in section 4, I will analyze in a forecasting experiment which forecasting model is able to approximate the respective survey closest. As suggested by the theory on heterogeneous expectations, aggregate measures of inflation expectations should be seen as a weighted average of different forecasting schemes. Hence, section 5 will be devoted to the questions which part of the survey participants learn by signal extraction and does the composition of aggregate expectations change when different periods are considered?

2 A First Look at Inflation Expectations

In the following section, some of the inherent features of inflation expectations and related forecast errors are explored for a selection of US surveys: the *Survey of Professional Forecasters* (SPF), the *Livingston Survey* (LIV) and the *Michigan Household Survey* (MHS). Thereby we will briefly test if the common notion of rational expectations in the way it has been introduced by Muth (1961) applies to survey measures of inflation expectations of experts.

2.1 Data Description

On the whole, I will focus on five different questions asked in the surveys mentioned above. The survey results, which will be labeled *SPF h=1* in the following contain the expected quarterly change of the GDP deflator one quarter ahead. Here, data is available from 1968 fourth quarter and ends in 2007 second quarter. *SPF h=4* gives information on the expected average change of the quarterly GDP deflator during the next four quarters. The dataset starts in 1970 second quarter and ends in 2007 second quarter. Note, that these forecasts are overlapping as the survey is conducted on a quarterly frequency. *LIV h=1* contains expectations of the annualized six month consumer price inflation six months ahead. This constitutes no overlapping forecasts as *LIV* is conducted biannually. In contrast to *LIV h=2*, which gives expectations of 12 month CPI inflation one year ahead and where the overlap is one period. Another survey measure of inflation expectations is given by *MHS h=12* where households are asked the following question:

> A: During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?

B: By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?

This entails an overlap of 11 periods. Note here, that the reference series are quite different for the respective survey not only as far as the measure of price increase – and the associated variability of the series – is concerned but also with respect to the forecasting horizon.

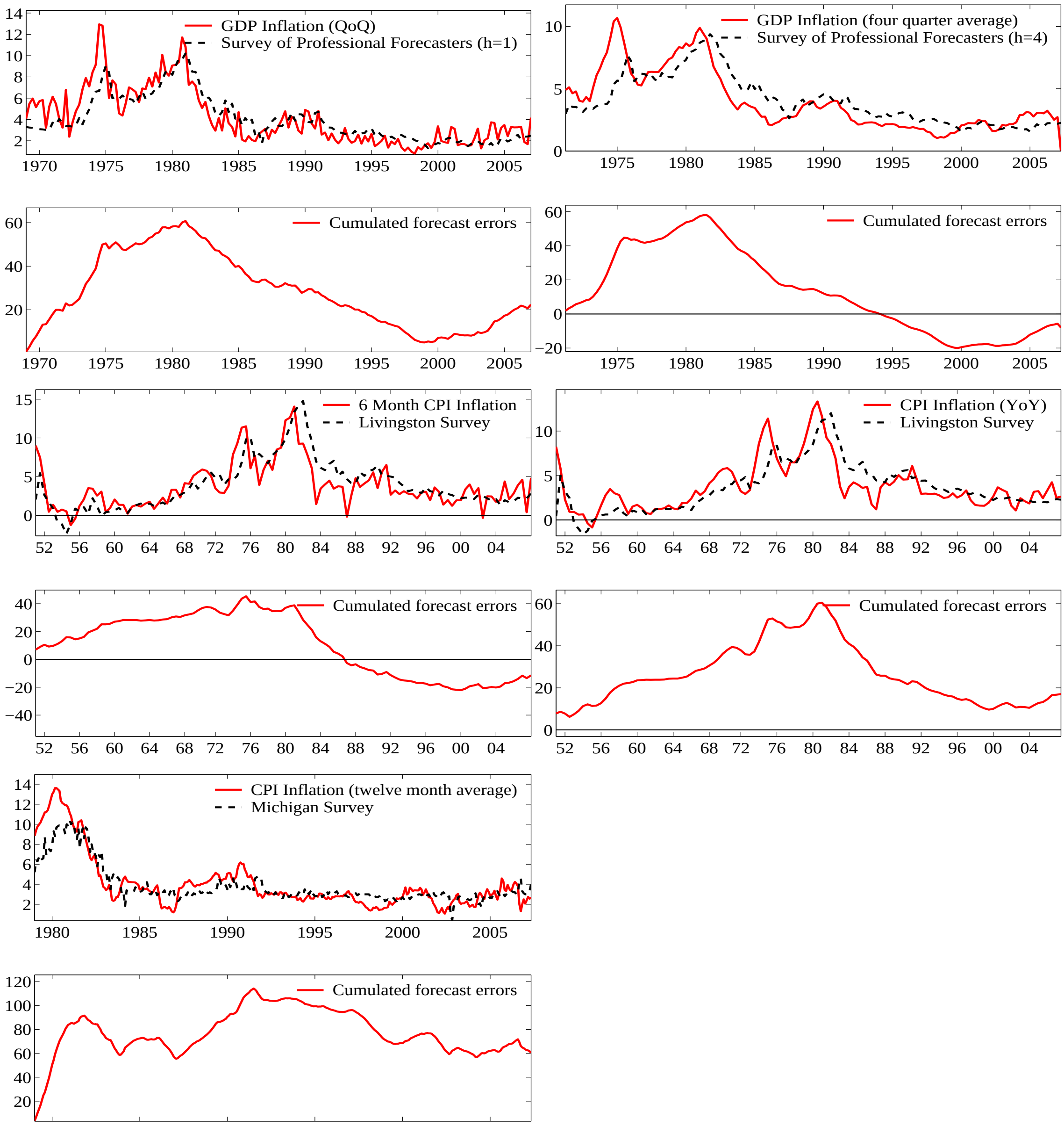
2.2 (Un)biasedness and (In)efficiency of Forecasts?

Define the survey expectation error as the difference between realized inflation and survey expectation with a forecast horizon of h periods. ($Error = \pi_t - \pi^e_{t|t-h}$). Thus, negative values result when the inflation rate is overestimated. Figure 1 visualizes the data by showing inflation expectations of *SPF*, *LIV* and *MHS* respectively. All surveys cover the period of high inflation beginning in the seventies, reaching a peak around 1980 and falling again in the subsequent period of disinflation under the Volcker regime. They also contain the rather tranquil period of the presidency of Greenspan starting in November 1987 and the recent period under the presidency of Bernanke since 2006. It becomes clear that cumulated forecast errors tend to follow the pattern of the inflation rate itself. This means that during phases of rising inflation like in the 1970s a repeated underprediction of inflation rates can be observed. As inflation comes down to moderate levels in due course, the cumulated forecast error decreases again most notably for *SPF* and *LIV*, which means that inflation is overpredicted during that period. Also note that in almost all cases considered here, the cumulated forecast error displays strong persistence. This means that an error in one period is not completely offset in the subsequent period but agents are sluggish when changing expectations. Thus, there seems to be a case for bounded rationality.

The recent findings can also be investigated more formally. In the following, I basically update some of the results on survey expectations found in Evans and Wachtel (1992) whose sample ends in 1991. Following the rational expectations hypothesis of Muth (1961), forecast errors as defined above should follow a zero mean white noise process if survey participants form rational expectations. This requires expectations to be unbiased and efficient in the sense that no information is omitted when forming expectations. To check if unbiasedness is a valid assumption, I run the regression described in equation (1) and test if $a = 0$ and $b = 1$ by means of a Wald test.

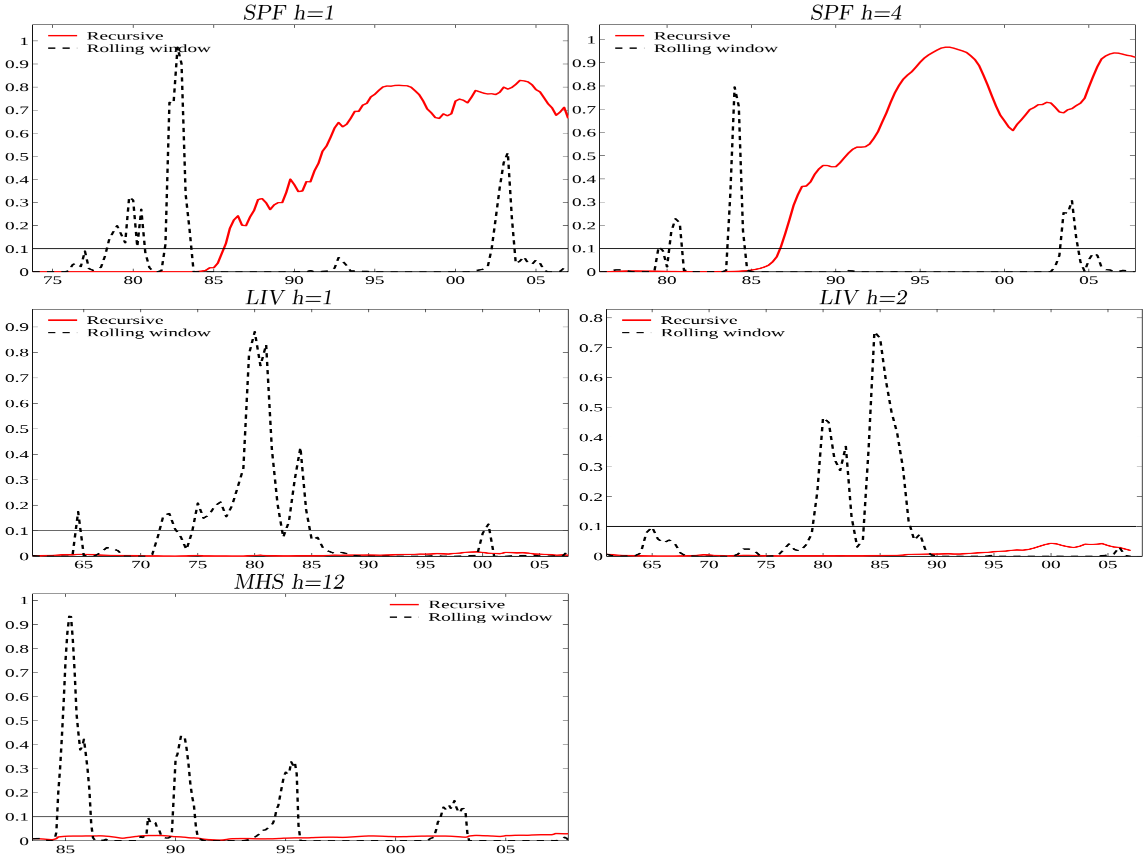
$$\pi_t = a + b\pi^e_{t|t-h} + \epsilon_t \tag{1}$$

Here, $\pi^e_{t|t-h}$ is the expected inflation rate conditional on the information set at time $t-h$. Figure 2 plots recursive Wald tests, as well as a test based on a rolling window of five years for *SPF* and *MHS* and ten years for *LIV*. Considering the whole sample, unbiasedness is not rejected for both questions asked in SPF. If sub–samples are considered by recursive estimation, SPF provides biased expectations up to the mid–eighties which coincides with the large swing of the inflation rate. Afterwards – when disinflation has come to an end – the test indicates that expectations are unbiased. Rolling window estimates point into the same direction. The finding can also be confirmed by looking at the cumulated forecast error which returns to zero in the mid–nineties, thereby indicating that – on average – expectations have been unbiased.



Note: The first panel shows the expected annualized quarterly GDP inflation one quarter ahead along with the realizations. The sample begins in 1968 Q4 and ends in 2007 Q2. The second panel depicts the four quarter moving average of GDP inflation along with expected average annualized inflation from *SPF* during the next $h = 4$ quarters where the sample runs from 1970 Q2 to 2007 Q2. The third panel depicts the annualized six months growth rate of CPI along with expected inflation with a forecasting horizon of $h = 1$ half years. The fourth panel contains the one year growth rate of prices along with expected inflation with a forecasting horizon of $h = 2$. Both measures are taken from *LIV* where the sample runs from 1950 I to 2007 I. The last panel shows CPI Inflation as the twelve months moving average growth rate of prices along with expected CPI inflation with a forecasting horizon of $h = 12$ months from *MHS*. The sample runs from 1978 M1 to 2007 M6. The lower part of each panel shows a plot of the cumulated forecast errors $\sum_{\tau=0}^{t}(\pi_\tau - \pi^e_{\tau|\tau-h})$ up to time t.

Figure 1: Inflation expectations from MHS, SPF, LIV



Note: The solid red line shows p–values for a recursive Wald–test of $H_0 : a = 0, b = 1$. The dashed line represents p–values based on a rolling window. The initial estimation period and the rolling window cover 5 years for *SPF*, 10 years for *LIV* and 5 years for *MHS*. The sample runs from 1968 Q4 - 2007 Q2 (*SPF h=1*), 1970 Q2 - 2007 Q2 (*SPF h=4*), 1950 I - 2007 I (*LIV h=1* and *LIV h=2*) and 1978 M1 - 2007 M6 (*MHS*).

Figure 2: Recursive Wald–test *SPF*, *LIV*, *MHS*

On the other hand, LIV, which is questioned on a semiannual frequency, is clearly biased. But when estimated on a rolling window beginning in the late seventies, which does not cover much of the period of high inflation, it turns out to be unbiased. The MHS is biased throughout the whole sample, whereas the rolling window tests indicate unbiasedness from time to time – especially during the mid–eighties again. Keeping in mind that it is an household survey and that the sample does not cover all of the high inflation period either, this does not come a surprise. When compared to the representation of the cumulated forecast error, biasedness is confirmed by the fact that the zero line is not crossed although the cumulated errors clearly stabilize in the second half of the sample. Thus, whether an expectations series is biased crucially hinges on the time period considered. One conclusion which can be drawn here is that biasedness of expectations seems to be a small–sample problem in the sense that samples are finite.

It is also worthwile to investigate if forecast errors are larger and tend to exhibit more persistence when the underlying variable experiences large changes. Therefore, the correlation between forecast errors and inflation is presented in the left part of table 1, whereas the right part gives the correlation with forecast changes.

	Cross correlation of forecast error							
	$\Delta\pi_t$				$\Delta\pi^e_{t\|t-h}$			
$\pi_t - \pi^e_{t\|t-h}$	lag_0	lag_1	lag_2	lag_4	lag_0	lag_1	lag_2	lag_4
SPF h=1	0.54	0.27	0.03	0.15	0.43	0.17	0.10	0.20
SPF h=4	0.73	0.72	0.66	0.44	0.76	0.78	0.74	0.51
LIV h=1	0.64	0.27	0.19	−0.04	0.55	0.20	0.14	−0.05
LIV h=2	0.77	0.59	0.33	−0.03	0.70	0.60	0.33	0.03
MHS h=12	0.76	0.73	0.66	0.56	0.61	0.55	0.53	0.51

Note: The sample of survey forecast errors runs from $1969Q1 - 2007Q2$ (*SPF h=1*), $1971Q2 - 2007Q2$ (*SPF h=4*) and $1950\,II - 2007\,I$ (*LIV h=1*) and $1951\,I - 2007\,I$ (*LIV h=2*) and $1979M1 - 2007M6$ (*MHS*). The displayed lag lengths coincide with very different time intervals. Due to the different frequencies of the surveys, four lags imply for the *MHS* 4 months, for the *SPF* one year and for the *LIV* two years.

Table 1: Cross correlation of forecast errors

Errors are apparently positively correlated with the change of the inflation rate. This is compatible with the view that an overestimation of inflation comes along whenever the inflation rate declines or has declined the period before. In other words, the higher the decline, the larger is the associated overestimation which implies that forecasters do not respond very rapidly to shocks in the inflation rate. Interestingly, the same result can also be found if we look at the correlation of forecast errors with forecast revisions. Whenever an underprediction occurs, there is a tendency to raise forecasts in subsequent periods. In general, this shows that forecasts do not respond very quickly to past errors.

To conclude, expectations are formed in a way inconsistent with the common concept of rationality which relies on full information and perfect foresight. Consequently, a number of studies have also come to the conclusion that rational expectations do not provide a good description of expectation formation processes[3]. It is important to note that expectational errors are found to be persistent especially in periods of large inflation movements. Moreover, for *SPF*, a bias is only found for such a period but not for the whole sample, whereas the other surveys are unbiased in a sub–sample around the mid–eighties. From table 1 it can also be inferred that forecast errors are larger during periods which are characterized by large swings of inflation and where forecasting is more complicated. Note, that all surveys seem to behave very similar with respect to bias and persistence of expectational errors despite the fact that respondents and reference variables as well as forecasting horizons differ considerably.

[3]See for example Roberts (1997) and the papers cited there.

3 Learning With a Simple Forecasting Model

3.1 Motivation and General Framework

Standard New Keynesian models, which assume rational expectations, will predict that unanticipated changes in the inflation target of a central bank will lead to an immediate jump of expectations and the level of inflation. As shown in section 2.2 this is not a realistic assumption because expectations are characterized by significant inherent persistence. Moreover, section 2.2 suggests that, although inflation may move rather quickly in disinflation episodes, expectations adjust only sluggishly. In Andolfatto and Gomme (2003), for instance, the authors argue that these features are observed not because agents are ignorant. They simply face a complicated forecasting problem which introduces sluggish adjustment to a new target inflation. Thus, disinflation comes along with significant output loss. The key assumption of the analysis in the present paper is that private decision makers cannot directly observe the inflation target pursued by the central bank. This may be due to the fact that the central bank is not transparent or that it has low credibility. Moreover, there may be information problems as far as the timing of the change is concerned. Note, that this situation is well applicable for the case of U.S. inflation during the eighties. Moreover, it may be valid even today as the FED does not announce an explicit target rate.

Following Andolfatto, Hendry, and Moran (2002), Dossche and Everaert (2005), Kozicki and Tinsley (2005) and Erceg and Levin (2003), I assume that agents solve a signal extraction problem in order to infer whether the observed movement of inflation is due to a transient monetary policy shock or whether the monetary authority has changed its inflation target, which will result in a permanent change of the level of inflation. In particular, they have to form expectations about trend inflation, which, in these models, coincides with the central bank's inflation target.

More intuitively, the signal extraction problem can be described as follows. In the first place, private decision makers observe an interest rate that is higher than would be the outcome of a strict application of the interest rule. They perceive the action of the central bank as if there had been a discretionary transitory shock to the system. However, it will be offset by application of the monetary policy rule in due course when the monetary authority wants to achieve its target again. Consequently, after having observed a monetary tightening, agents expect that, in the next quarter, inflation will rise to target levels again. If, however, inflation remains at lower than target levels for more a longer period of time, then decision makers will conclude that the target has changed. But they are not informed about the magnitude of the change. This constitutes the signal extraction problem as they cannot observe to which extent the movement of the inflation rate is due to a temporary monetary action and which part of the change in inflation rates is the result of a permanent shift. Thus, they have to form expectations about the target that is pursued by the central bank. Even if they do so in an optimal way, it will take some time until they can infer the correct new target from the observed action. Consequently, expectational errors will occur although decision makers are not ignorant. A fact that has been emphasized by Andolfatto, Hendry, and Moran (2006) and Evans and Wachtel (1992).

One obvious solution to the signal extraction problem is given by application of the Kalman filter. Due to its recursive formulation, it provides the theoretical concept describing how private agents learn about trend inflation in this study. Moreover, this learning rule is optimal, provided that agents know the true model of the economy. However, as for instance argued in Branch and Evans (2006), the concept of bounded rationality is more appropriate here. According to the concept, agents do not know the true model of the economy in reality but they will rather apply a simple model that is easily applicable and serves the respective purpose. For instance, resource constraints or limited tools for information processing will advocate the use of a simple model on behalf of private agents.

Moreover, agents may be quite heterogeneous with respect to resource constraints or as far as the evaluation of a bad forecast is concerned. Therefore, the theory of heterogeneous expectations outlined in – among others – Branch (2004) suggests that each agent undertakes a different effort to find a suitable forecasting model and to get an estimate of the model's parameters. That is one reason why one part of the agents may, for instance, be rational or trend learner, another part will simply be backward–looking or completely ignorant.

This is the framework for the present analysis, where, on the one hand, survey participants are characterized by a lack of information with respect to target inflation. On the other hand, they will rely on a simple forecasting model. In particular, they do not make the effort to find the true model of the economy. Moreover, it may be the case that more than one "type" of forecaster contributes to the survey result.

3.2 The Forecasting Model

In the following, I will formalize expectation formation of private agents. It will become clear that an unobserved components model is very well suited to model the framework outlined in section 3.1. In the spirit of bounded rationality, I assume that agents form expectations according to a simple forecasting model. The basic model consists of three (unobserved) components. It comprises a time–varying trend $\overline{\pi}_t$ that captures the permanent component of inflation and, in addition, inflation inherits cyclical movements $\widehat{\pi}_t$ and unsystematic shocks ϵ_t. The model is given by equations (2) to (4) which constitute the data generating process for inflation expectations.

$$
\begin{aligned}
&(2) \qquad & \pi_t &= \overline{\pi}_t + \widehat{\pi}_t + \beta d_t + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \\
&(3) \qquad & \overline{\pi}_{t+1} &= \overline{\pi}_t + \eta_t \qquad \eta_t \sim N(0, \sigma_\eta^2) \\
&(4) \qquad & \begin{pmatrix} \widehat{\pi}_{t+1} \\ \widehat{\pi}^*_{t+1} \end{pmatrix} &= \rho \begin{pmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{pmatrix} \begin{pmatrix} \widehat{\pi}_t \\ \widehat{\pi}^*_t \end{pmatrix} + \begin{pmatrix} \kappa_{t+1} \\ \kappa^*_{t+1} \end{pmatrix} \\
& & \begin{pmatrix} \kappa_{t+1} \\ \kappa^*_{t+1} \end{pmatrix} &\sim N(0, \begin{pmatrix} \sigma_\kappa^2 & 0 \\ 0 & \sigma_\kappa^{2*} \end{pmatrix})
\end{aligned}
$$

This is basically the widely used *trend plus cycle model* introduced by Harvey (1989) augmented by a vector of dummy variables d_t. Here, the trend is assumed to follow a random walk and the transitory component can be shown to follow a stationary ARMA(2,1) process with complex roots. Above all, the model accounts for the fact that there is a signal extraction problem as the distinct components are not observable but have to be estimated. When learning about the unobserved target, forecasters are assumed to update their perception each time a new observation becomes available. To be more precise, they learn from their forecast errors of the past. Like in Erceg and Levin (2003) agents make use of a so–called constant gain learning rule. This means that, after having observed a forecast error, each component is updated by a constant part of the error. Technically speaking, when estimating the unobserved components, agents learn from noisy information contained in the forecast errors $\nu_t = \pi_t - E_{t-h}\pi_t$. In particular, there is learning of both, trends as well as cycles at the same time. This is in contrast to the standard adaptive learning scheme like in Pajfar and Santoro (2007), for example. Consequently, forecasters will behave very much like econometricians who estimate an unobserved components model. In order to solve the signal extraction problem optimally, they can make use of the Kalman filter recursions to obtain an estimate of the unobserved components. The optimal gain is then given by the so–called Kalman gain.[4] Note, that in the event of changing trend inflation expectations would necessarily be biased and forecast errors are persistent, a feature we also find for inflation expectations during the Volcker era in section 2.2.

[4]See Harvey (1989), chapter 3.2.

3.3 Fitting Survey Expectations

3.3.1 The Model

In the following, an in–sample perspective is taken. If the data generating process can be described by equations (2) to (4), the process for inflation expectations can be written down in the form of a state–space model which consists of an observation equation (5) and state equations (6) to (8), which describe how unobserved components are estimated. Thus, a forecast for π is calculated from

$$\pi_{t+1|t} = \overline{\pi}_{t+1|t} + \widehat{\pi}_{t+1|t} + \beta d_{t+1}. \tag{5}$$

The subscript $t|t-1$ denotes the mean of the distribution at t predicted from information up to time $t-1$. The Kalman filter recursions, which are employed to estimate the unobserved components are reformulated such that they only contain predicted state variables. Expectations of trend and cycle in the next period (based on the last prediction) are then given by:

$$\begin{aligned}
\overline{\pi}_{t+1|t} &= \overline{\pi}_{t|t-1} + \mathcal{K}_{1,t}\nu_t &(6)\\
\widehat{\pi}_{t+1|t} &= \rho\cos\lambda\widehat{\pi}_{t|t-1} + \rho\sin\lambda\widehat{\pi}^*_{t|t-1} + \widetilde{\mathcal{K}}_{2,t}\nu_t &(7)\\
\widehat{\pi}^*_{t+1|t} &= -\rho\sin\lambda\widehat{\pi}_{t|t-1} + \rho\cos\lambda\widehat{\pi}^*_{t|t-1} + \widetilde{\mathcal{K}}_{3,t}\nu_t &(8)
\end{aligned}$$

Here, $\nu_t = \pi_t - \pi_{t|t-1}$ denotes the expectation error of the last period. It can be shown that $\widetilde{\mathcal{K}}_{2,t} = \rho\cos\lambda\mathcal{K}_{2,t} + \rho\sin\lambda\mathcal{K}_{3,t}$ and $\widetilde{\mathcal{K}}_{3,t} = -\rho\sin\lambda\mathcal{K}_{2,t} + \rho\cos\lambda\mathcal{K}_{3,t}$ where $\mathcal{K}_{i,t}$ represents the gain parameter according to which unobserved components are updated when a misperception of inflation occurs[5]. In particular, $\mathcal{K}_{1,t}$ determines the update of the estimated trend and $\mathcal{K}_{2,t}$ determines the updating scheme with respect to the transitory part. $\mathcal{K}_{3,t}$ captures an indirect effect of misperceptions on the update of the transitory component and is given for completeness. The optimal forecasting scheme with respect to the data generating process given in section 3.2 is given by the Kalman filtering rule. Thus, trend expectations should be updated by an amount equal to the implied Kalman gain.

In order to investigate the properties of the survey, the conditional mean in the future $\pi_{t+h|t}$ is replaced by survey expectations $\pi^e_{t+h|t}$. The forecast error ν_t is exchanged with its observed counterpart, associated with the respective survey with forecast horizon h, i.e. $\nu^e_t = \pi_t - \pi^e_{t|t-h}$. In the cases where $h > 1$, this would

[5]Note that gain parameters $\mathcal{K}_{i,t}$ relate to reduced form parameters $\widetilde{\mathcal{K}}_{i,t}$ in a linear way.

imply that forecasters apply some kind of direct multi–step forecasting. Hence, the gain parameters cannot be interpreted like the usual Kalman gains. In other words, signal extraction with the Kalman filter will only yield a minimum forecast error, if it relies on the one–step–ahead forecast error. Hence, in–sample results are presented for *SPF h=1* and *LIV h=1* only. Finally, ε_t reflects the part of survey expectations which is not explained by the model.

In this simple univariate setting, equations (7) and (8) capture the persistence of the transitory part of expectations. In addition, the model allows for signal extraction – i.e. learning from repeated forecast errors, because trend learning is necessarily associated to the estimation of the cyclical component. Hence, a part of the forecast error is related to misperceptions of the cyclical part. Thus, we obtain in–sample estimates for the gain parameters $\mathcal{K}_{1,t}$ to $\mathcal{K}_{3,t}$ for each survey. These determine the speed of learning of survey participants when a change of unobserved trend inflation occurs. Furthermore, we can test if extracted trend expectations are characterized by persistence and it is possible to infer the speed of trend learning.

3.3.2 Estimation Results

The system which consists of equations (5) to (8) is estimated by maximum likelihood. The diffuse likelihood is computed by the Kalman filter with diffuse prior density of the initial state vector. The parameter vector $\psi = \begin{pmatrix}\sigma^2_\varepsilon & \lambda & \rho\end{pmatrix}$ consists of the variance of the irregular component ε_t, the cycle length λ and the so–called dampening factor ρ. It is reparameterized such that the theoretical restrictions are fulfilled (see appendix A for details). Dummy variables are set when the outlier test proposed by Harvey and Koopman (1992) indicated an outlier. Common regression diagnostics and a histogram of past forecast errors ν_t are given in appendix C. The constant gain parameters $\mathcal{K}_i$ can be extracted from the smoothed state vector and are not restricted during estimation, whereas the smoothing recursions also yield estimated standard errors. The estimated parameters are summarized in table 2. The in–sample observation period runs from 1972 to 2007. Hence, there is at least one possible structural break agents may have learned which is commonly associated with the beginning of the Volcker era. Taking a look at equations (5) to (8), it becomes clear that there is only one error term in the system (ε_t) which captures irregularities. Hence, in a technical sense, the estimated unobserved components are non–stochastic as far as the Kalman recursions are concerned – i.e. the dynamics of all dependent variables is solely explained by past forecast errors and autoregressive elements.

Turning to table 2, one observes that estimated variances σ^2_ε, which are presented in the first column, are in a plausible range and significant at common levels. In the case of *SPF h=1* the estimated cycle parameter $\widehat{\lambda}$ is 0.26 which implies a cycle

		$\widehat{\sigma}_\varepsilon^2$	$\widehat{\lambda}$	$\widehat{\rho}$	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$
SPF h=1	72–07	0.469	0.261	0.953	0.119	0.010	0.066
		(0.058)	(0.021)	(0.012)	(0.004)	(0.018)	(0.013)
LIV h=1	72–07	2.020	2.307	0.000	0.136	0.016	0.063
		(0.389)	(–)	(–)	(0.009)	(–)	(–)

Note: The table shows maximum likelihood estimates of the parameters and estimated constant gain coefficients of the surveys. Numbers in parenthesis are standard errors. A (–) indicates that numerical estimates are not available, which is the case if ρ hits the lower bound of zero. Estimated parameters are presented for completeness.

Table 2: Estimated parameters

length of approximately 24 quarters. $\widehat{\rho}$ which determines the "sluggishness" of the cycle is in a plausible range being compatible with the concept of an autoregressive transitory component. As far as *LIV h=1* is concerned, $\widehat{\rho}$ hits the lower bound of zero, which means that the cyclical element is not existent. As the frequency of the survey is only biannual, it is likely that transitory movements of inflation expectations display a rather irregular pattern over the business cycle. Hence, much of the variation of expectations in *LIV h=1* is captured by the error term, which has a rather high variance. Most notably, the model seems to be supportive for the trend learning hypothesis. The gain parameters of *SPF* and *LIV*, although not restricted during estimation, lie between 0.12 and 0.14, depending on the survey. This means that in the case of *SPF h=1*, for instance, about 12 percent of the last error is attributable to trend mis–perceptions. Moreover, in both surveys, $\mathcal{K}_1$ is highly significant which leads to the conclusion that trend updating is an important characteristic of inflation expectations.

3.3.3 Trend and Cyclical Components of Expectations

Figures 3 to 4 depict the unobserved components which are extracted by the Kalman smoothing recursions beginning in 1972.

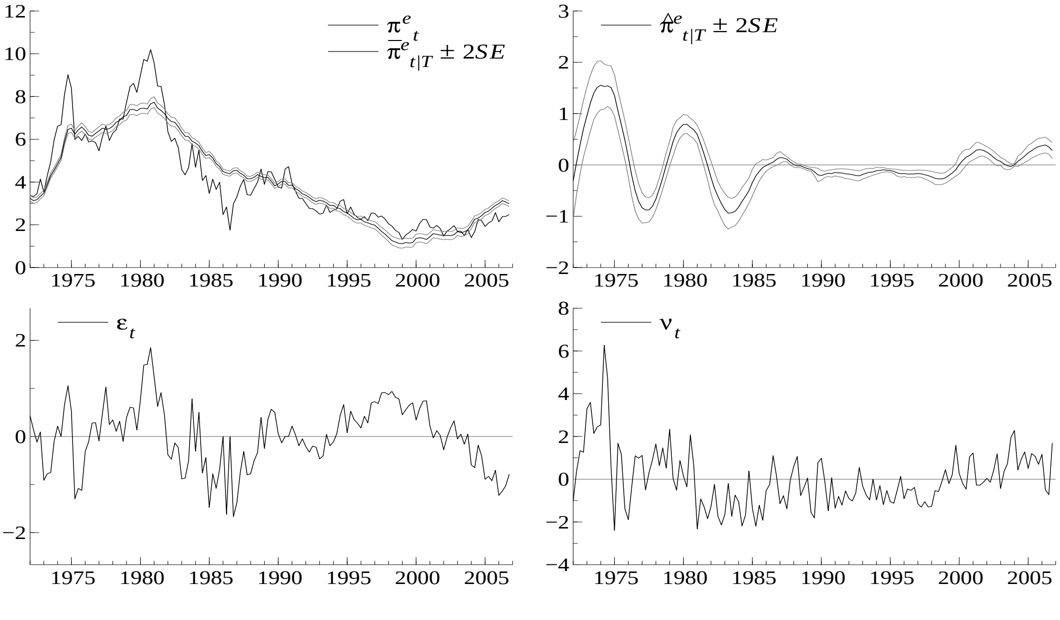
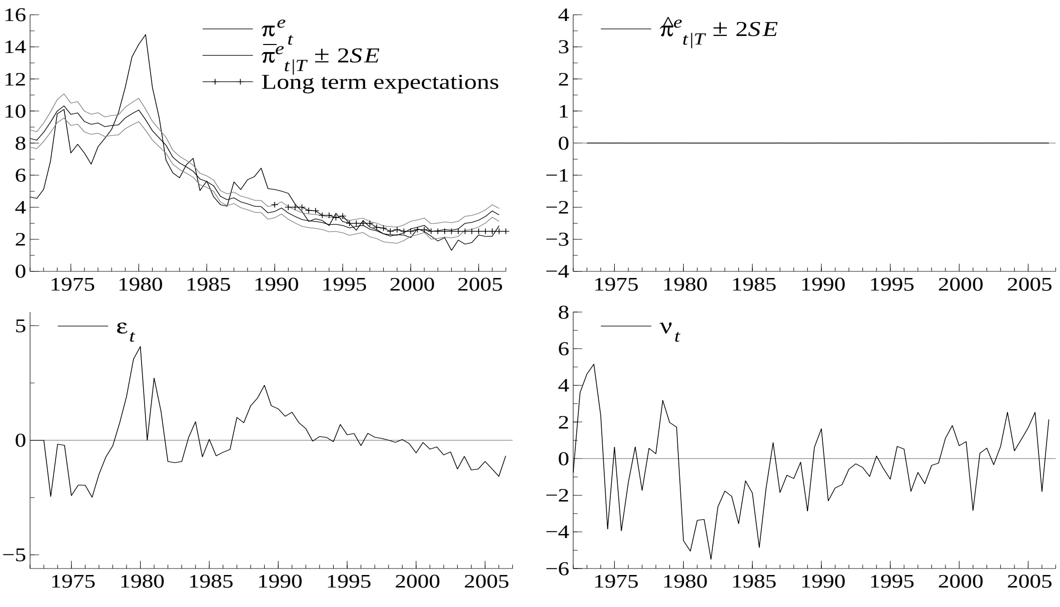


Figure 3: Learning model *SPF h=1*



Note: The upper left panel shows inflation expectations together with the smoothed expected trend $\pi^e_{t|T}$ where T is the last available observation. The second panel depicts the smoothed cyclical component $\widehat{\pi}^e_{t|T}$. The irregular component ε_t and the forecast error ν_t are given in the lower part.

Figure 4: Learning model *LIV h=1*

The upper left hand graph depicts the original expectation series π_t^e together with the estimated trend component and an error band of two standard deviations. In general, it becomes apparent that trend expectations do not jump. By contrast, they are very sluggish and move rather slow. Note that, although stemming from different surveys and comprising quite different target variables, the general picture presented in both figures is quite similar. As a consequence, it takes until the mid–nineties to obtain trend expectations that are around 3%. Turning to *SPF*, this is reflected by the fact that estimated gain parameters $\mathcal{K}_1$ in table 2 are rather low. This implies that each quarter – by and large – only ten percent of the forecast error is attributable to trend misperceptions.

Inspection of the transient component reveals that cyclical movements are much more pronounced until the mid–eighties. The same seems to be true for forecast errors. As the model is written, the irregular component of the signal equation captures all of the unexplained part of the dynamics. Secondly, however, there may be a time–varying nature to the expectations formation process. It might be the case that during tranquil periods, like after 1987, it may be worthwhile for the agents to adapt a simple backward–looking forecasting scheme. This could, in principal be tested by splitting the sample. However, for an estimation of the structural time–series model the period is rather short. But also note that we observe a change in the behavior of ε_t, which seems to display less systematical movements during the Volcker period. That is one reason why, for the later analysis, the sample will be split in 1987 (see section 4.3 and 5).

Now turning to *LIV h=1*, we observe a flat cyclical pattern, whereas the residual component captures most of the dynamics. Hence, past forecast errors seem to explain trend dynamics but do not account for the transitory movements. However, the model seems not to be completely at odds, as it is capable to reproduce a trend that is consistent with long–term inflation expectations contained in the Livingston Survey. Moreover, the trend is learned quite slowly as estimated gain parameters imply that every half year by and large 15% of the error is used to adjust trend expectations.

Generally speaking, it is possible to fit survey data on inflation expectations to the simple learning model presented here. It produces the sluggishness of expectations in the event of shifts in target inflation. The reason is that agents are learning from a noisy signal which – in the univariate setup here – is the past forecast error. Note, that $\mathcal{K}_i$ has not been restricted during estimation. Nevertheless, $\mathcal{K}_1$ has the correct sign and is significant, meaning that an underprediction leads to an upward revision of the trend. As far as the cyclical component is concerned, results are less clear–cut. This is due to the fact that I employ the same model for both survey measures. As there are differences with respect to the underlying frequency of the surveys, very different cyclical patterns may emerge, moreover results for *LIV h=1* rely on less observations. Another implication of the present findings is, that there may be some time–variation of expectation formation schemes depending on the distinct presidential periods. One might conjecture that for the Volcker disinflation signal extraction seems to work better than for the later period.

4 Forecasting Inflation

4.1 The Simulation Experiment

In the following, I will simulate the forecasting exercise undertaken by survey participants taking an out–of–sample perspective. Equations (2) to (4) constitute the data generating process for inflation expectations, where the forecaster is assumed to behave like an econometrician. Similar to Branch and Evans (2006), the forecasting procedure can then be split up into three steps. In a first period, the forecaster gains some experience over the dynamics of inflation rates. In this in–sample period, he estimates the parameters of the model and runs the Kalman recursions to obtain estimates of the unobserved components. He also observes the updating gain implied by the Kalman filter. This in–sample estimation period starts at the beginning of 1953 and ends in 1980 for all models. Thus, the first period covers 27 years of data which should suffice to shape the experience of a forecaster – i.e. to obtain reliable estimates. In a second step, the forecaster takes the estimated parameters as given and extracts the unobserved components up to the last published record of inflation by relying on the gain parameters estimated during the first period. This is done subsequently for each observation following this in–sample period. In the third step, the forecaster then generates an out–of–sample forecast of the signal variable – i.e. inflation. The forecast horizon is chosen such that it matches the respective survey forecast.[6] Note that a number of different models have to be built, as survey expectations involve different target variables. In addition, some exogenous variables have been added to the forecasting models to account for the fact that survey participants might also look at aggregate output or interest rates when forming their forecasts. As benchmark cases, I also introduce a naive forecasting scheme (Model I) and a simple AR(1) model in first differences (Model II). The reason is, that the model does not involve trend considerations explicitly. Nevertheless, some learning takes place because parameters are estimated by recursive least squares. To be more precise, Model II comes very close to the type of learning models employed by – among others – Branch and Evans (2006). It has the features of a widely im-

[6] The forecast horizon is the next quarter (*SPF h=1*), the average of the next 4 quarters (*SPF h=4*), the next half year (*LIV h=1*), the next year (*LIV h=2*) and the average of the next 12 months (*MHS* h=12).

plemented (decreasing gain) learning algorithm as parameter estimates are updated every time a new observation becomes available[7]. I also introduce perfect foresight, which provides rational expectations (Model VII). In detail, the following models have been employed:

Model I: $E_t\pi_{t+h} = \pi_t$

Model II: $\Delta\pi_t = \alpha_0 + \alpha_1\Delta\pi_{t-1} + \varepsilon_t$ estimated recursively

Model III: $\pi_t = \overline{\pi}_t + \widehat{\pi}_t + d_t + \varepsilon_t$

Model IV: $\pi_t = \overline{\pi}_t + \widehat{\pi}_t + d_t + \sum_{\tau=0}^{3} \Delta i_{t-h-\tau} + \sum_{\tau=0}^{3} \Delta y_{t-h-\tau} + \varepsilon_t$

Model V: Model III estimated recursively

Model VI: Model IV estimated recursively

Model VII: $E_t\pi_{t+h} = \pi_{t+h}$

Here, Δi_t denotes the change of the three month treasury bill rate measured either on a quarterly frequency (*SPF*) or on a monthly frequency (*LIV, MHS*). Δy_t is a measure of aggregate output growth, which means that the change of industrial production has been employed on a monthly frequency and GDP growth for *SPF*, which is observed on a quarterly frequency.[8]

Models III and VI are estimated by maximum likelihood, whereas the likelihood is computed by the Kalman filter, which is initialized with a diffuse prior density of the state vector. The parameter vector, here, comprises five variables $\psi = \begin{pmatrix}\sigma_\eta^2 & \sigma_\kappa^2 & \lambda & \rho & \sigma_\varepsilon^2\end{pmatrix}$. During estimation ψ is reparameterized according to theoretical restrictions (see appendix B for details). Models V and VI are essentially the same as Models II and III but estimated recursively. This means that forecasters come up with a new set of maximum likelihood estimates whenever new data becomes available. Most importantly, also the estimated gain parameters will change with every new observation. This is a more sophisticated forecasting scheme than before, as it assumes that forecasters revise their estimates from time to time. But it also represents a learning scheme where agents learn from past misperceptions.

The different forecasting models are now used to forecast five different target series, which are chosen such that they match the respective survey (see figure 1 for

[7]See Evans and Honkapohja (2001) for further details on recursive least squares learning schemes. Here, it is generally assumed that private agents learn the parameter values of the rational expectations solution of the model. Also see Branch and Evans (2006) and Weber (2007) and the papers cited there for empirical approaches.

[8]Note that π_t is the quarterly change of GDP inflation for *SPF h=1*, the average annualized GDP inflation during the next four quarters for *SPF h=4*, the annualized 6 months CPI inflation for *LIV h=1*, the twelve months CPI inflation for *LIV h=2* and the average annualized CPI inflation during the next twelve months for *MHS h=12*.

a graphical representation of the time series to be forecast). Note, that simulated forecasters use monthly data for six and twelve month CPI inflation but the forecast is made on a semi–annual frequency. However, for the twelve month average CPI inflation series, the forecast is produced every month. To ensure that the simulated forecasters start out with a well specified model in 1980, some dummy variables are introduced to the model whenever the outlier t–test proposed by Harvey and Koopman (1992) showed signs of severe outliers.[9]

4.2 Estimated Parameters

Next, we take a closer look at recursively estimated parameters from Models V and VI to see whether they change over time. The recursive estimates of the hyperparameters as well as the implied (optimal) gain $\mathcal{K}_1$ to $\mathcal{K}_3$ are depicted in appendix D in figures 6 to 9.[10] On the whole, hyperparameters are stable; only estimated variances seem to display some tendency to fall over time. In the one or the other case, the simulation exercise converges to a solution that implies a jump in parameter estimates, which would then show up as a distinct peak or drop in the series. Here, these occurrences are simply taken as given and can be interpreted as the difficulty of the respective forecaster in finding an appropriate model at each point in time. Turning now to the case of GDP inflation, $\mathcal{K}_1$ starts with values around 0.40 and slowly rises to 0.50 until around 1987. A very similar pattern emerges for the six month CPI inflation model. The twelve month CPI inflation model updates trend forecasts with a gain parameter which implies that about 70% of the error is associated with trend misperceptions. Hence, there seems to be a minor variation in the (optimal) learning rule for periods of large trend shifts.

Next, we can compare the gain parameters which apply for the out–of–sample models with those that have been estimated for the surveys in–sample. It becomes apparent that in an out–of–sample experiment forecasters would change their trend perceptions much more often. Results for *SPF h=1* imply that agents could improve their forecast performance by putting more weight on trend shifts – i.e. increasing the gain parameter $\mathcal{K}_1$ from 0.12 to about 0.50. Note, that semi–annual CPI forecasts with a 6 month horizon are based on a monthly model, which allows for a trend update every month. This is not the case for the corresponding survey (*LIV h=1*) in

[9] A couple of diagnostics tests have been conducted for each model. On the whole, the tests give satisfying results and indicate that the forecasting exercise relies on well specified models, although, for each target variable, the same forecasting model has been used. The results and estimates for the smoothed components are available upon request.

[10] For the presentation of gain parameters, it is generally assumed that, once having estimated the hyperparameters, the covariance matrix of innovations converges to the steady–state solution when the Kalman filter is run up to the last observation. Hence, the graphs show the estimated gain parameters conditional on the whole data set available at the time the forecast is made.

the in–sample analysis, where a trend update is based on semi–annual observations. Consequently, estimates for *LIV h=1* cannot be compared directly to the out–of–sample results of the present section.

4.3 Approximation of Survey Expectations

It is now important to see whether the simulated forecast series $\pi^f_{t+h|t}$ match the survey–based measures $\pi^e_{t+h|t}$. The approximation properties can be tested by the augmented version of the Diebold–Mariano statistic[11]. The resulting test statistics can be inferred from tables 3 to 5. In addition, the sample is split into two parts, the first one covering the whole sample 1980–2007, the middle panel covers the Volcker disinflation period 1980–1987 and the last panel covers the moderate period 1988–2007.

[11] Here, the null hypothesis is given by $H_0 : E[|\pi^e_{t+h} - \pi^{f,i}_{t+h|t}| - |\pi^e_{t+h} - \pi^{f,j}_{t+h|t}|] = 0$.

Model	SPF h=1 80–07 I	II	III	IV	V	VI	VII	SPF h=4 80–07 I	II	III	IV	V	VI	VII
I	0.00	5.84	6.45	4.95	6.37	4.42	-0.61	0.00	-2.21	0.49	-1.41	0.00	-0.83	-1.98
II	-5.84	0.00	2.04	2.05	1.67	1.41	-2.44	2.21	0.00	1.51	-0.94	1.26	0.28	-1.67
III	-6.45	-2.04	0.00	1.10	-2.20	0.10	-3.16	-0.49	-1.51	0.00	-2.04	-2.13	-1.33	-2.15
IV	-4.95	-2.05	-1.10	0.00	-1.62	-2.72	-3.34	1.41	0.94	2.04	0.00	1.79	1.57	-1.09
V	-6.37	-1.67	2.20	1.62	0.00	0.64	-2.95	0.00	-1.26	2.13	-1.79	0.00	-1.01	-1.97
VI	-4.42	-1.41	-0.10	2.72	-0.64	0.00	-2.94	0.83	-0.28	1.33	-1.57	1.01	0.00	-1.78
VII	0.61	2.44	3.16	3.34	2.95	2.94	0.00	1.98	1.67	2.15	1.09	1.97	1.78	0.00

Model	SPF h=1 80–87 I	II	III	IV	V	VI	VII	SPF h=4 80–87 I	II	III	IV	V	VI	VII
I	0.00	3.33	3.56	2.27	3.12	1.82	-0.80	0.00	-1.85	1.15	-0.93	0.72	-0.39	-1.47
II	-3.33	0.00	1.47	0.95	0.69	0.47	-1.91	1.85	0.00	2.09	-0.61	1.72	0.40	-1.25
III	-3.56	-1.47	0.00	0.15	-4.49	-0.53	-2.68	-1.15	-2.09	0.00	-2.24	-2.78	-1.47	-2.23
IV	-2.27	-0.95	-0.15	0.00	-0.82	-2.04	-2.34	0.93	0.61	2.24	0.00	1.79	1.15	-1.00
V	-3.12	-0.69	4.49	0.82	0.00	0.15	-2.23	-0.72	-1.72	2.78	-1.79	0.00	-1.10	-1.90
VI	-1.82	-0.47	0.53	2.04	-0.15	0.00	-1.95	0.39	-0.40	1.47	-1.15	1.10	0.00	-1.52
VII	0.80	1.91	2.68	2.34	2.23	1.95	0.00	1.47	1.25	2.23	1.00	1.90	1.52	0.00

Model	SPF h=1 88–07 I	II	III	IV	V	VI	VII	SPF h=4 88-07 I	II	III	IV	V	VI	VII
I	0.00	4.83	6.38	6.12	6.27	5.99	-0.08	0.00	-1.43	-0.71	-1.23	-1.00	-0.93	-1.63
II	-4.83	0.00	1.59	3.43	2.53	3.36	-1.56	1.43	0.00	0.13	-0.77	-0.05	-0.17	-1.23
III	-6.38	-1.59	0.00	3.14	0.81	1.91	-1.88	0.71	-0.13	0.00	-0.99	-0.78	-0.30	-1.29
IV	-6.12	-3.43	-3.14	0.00	-2.89	-2.00	-2.37	1.23	0.77	0.99	0.00	0.89	3.19	-0.50
V	-6.27	-2.53	-0.81	2.89	0.00	1.99	-1.95	1.00	0.05	0.78	-0.89	0.00	-0.12	-1.25
VI	-5.99	-3.36	-1.91	2.00	-1.99	0.00	-2.18	0.93	0.17	0.30	-3.19	0.12	0.00	-1.07
VII	0.08	1.56	1.88	2.37	1.95	2.18	0.00	1.63	1.23	1.29	0.50	1.25	1.07	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 108$ is the number of out–of–sample forecasts for the full sample. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 1.98 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 3: Modified Diebold–Mariano test on deviations of *SPF* and forecasting models

	LIV h=1 80–07							LIV h=2 80–07						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	0.17	0.76	0.34	1.50	0.93	-1.43	0.00	0.43	-0.30	-0.24	-0.52	-0.36	-1.20
II	-0.17	0.00	0.71	0.28	1.46	0.86	-1.39	-0.43	0.00	-0.34	-0.28	-0.55	-0.39	-1.22
III	-0.76	-0.71	0.00	-0.50	3.54	0.47	-1.52	0.30	0.34	0.00	0.42	-1.04	-0.15	-1.13
IV	-0.34	-0.28	0.50	0.00	1.38	3.36	-1.47	0.24	0.28	-0.42	0.00	-1.03	-0.61	-1.12
V	-1.50	-1.46	-3.54	-1.38	0.00	-0.54	-1.81	0.52	0.55	1.04	1.03	0.00	0.60	-0.94
VI	-0.93	-0.86	-0.47	-3.36	0.54	0.00	-1.71	0.36	0.39	0.15	0.61	-0.60	0.00	-1.07
VII	1.43	1.39	1.52	1.47	1.81	1.71	0.00	1.20	1.22	1.13	1.12	0.94	1.07	0.00
	LIV h=1 80-87							LIV h=2 80-87						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	0.65	1.49	1.10	1.59	1.16	-1.24	0.00	0.52	0.12	0.24	-0.09	0.04	-0.89
II	-0.65	0.00	1.31	0.90	1.42	0.97	-1.26	-0.52	0.00	0.06	0.18	-0.14	-0.01	-0.92
III	-1.49	-1.31	0.00	-0.16	0.98	-0.03	-1.72	-0.12	-0.06	0.00	0.93	-1.31	-0.37	-1.28
IV	-1.10	-0.90	0.16	0.00	0.32	0.67	-1.84	-0.24	-0.18	-0.93	0.00	-1.52	-1.57	-1.30
V	-1.59	-1.42	-0.98	-0.32	0.00	-0.21	-1.78	0.09	0.14	1.31	1.52	0.00	0.46	-1.02
VI	-1.16	-0.97	0.03	-0.67	0.21	0.00	-1.80	-0.04	0.01	0.37	1.57	-0.46	0.00	-1.12
VII	1.24	1.26	1.72	1.84	1.78	1.80	0.00	0.89	0.92	1.28	1.30	1.02	1.12	0.00
	LIV h=1 88–07							LIV h=2 88–07						
Model	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
I	0.00	-0.65	-1.24	-1.70	0.09	-0.54	-0.69	0.00	0.00	-0.98	-1.20	-0.95	-0.96	-0.90
II	0.65	0.00	-0.96	-1.38	0.38	-0.23	-0.61	0.00	0.00	-0.93	-1.17	-0.93	-0.92	-0.87
III	1.24	0.96	0.00	-1.22	3.79	1.44	-0.29	0.98	0.93	0.00	-0.77	0.13	0.48	-0.04
IV	1.70	1.38	1.22	0.00	3.35	4.01	-0.16	1.20	1.17	0.77	0.00	1.14	1.93	0.06
V	-0.09	-0.38	-3.79	-3.35	0.00	-1.25	-0.66	0.95	0.93	-0.13	-1.14	0.00	0.37	-0.06
VI	0.54	0.23	-1.44	-4.01	1.25	0.00	-0.51	0.96	0.92	-0.48	-1.93	-0.37	0.00	-0.12
VII	0.69	0.61	0.29	0.16	0.66	0.51	0.00	0.90	0.87	0.04	-0.06	0.06	0.12	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 54$ is the number of out–of–sample forecasts. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 2.00 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 4: Modified Diebold–Mariano test on deviations of *LIV* and forecasting models

	MHS h=12 80–07						
Model	I	II	III	IV	V	VI	VII
I	0.00	-2.79	-0.99	-0.99	-0.98	0.04	-0.16
II	2.79	0.00	-0.74	-0.73	-0.73	0.44	0.21
III	0.99	0.74	0.00	0.41	0.26	1.31	0.92
IV	0.99	0.73	-0.41	0.00	-0.04	1.34	0.93
V	0.98	0.73	-0.26	0.04	0.00	1.30	0.90
VI	-0.04	-0.44	-1.31	-1.34	-1.30	0.00	-0.25
VII	0.16	-0.21	-0.92	-0.93	-0.90	0.25	0.00
	MHS h=12 80-87						
Model	I	II	III	IV	V	VI	VII
I	0.00	-2.30	-0.87	-0.80	-0.96	0.44	0.75
II	2.30	0.00	-0.70	-0.63	-0.77	0.67	0.91
III	0.87	0.70	0.00	0.95	-0.32	1.69	1.68
IV	0.80	0.63	-0.95	0.00	-0.74	1.61	1.62
V	0.96	0.77	0.32	0.74	0.00	1.99	1.88
VI	-0.44	-0.67	-1.69	-1.61	-1.99	0.00	0.61
VII	-0.75	-0.91	-1.68	-1.62	-1.88	-0.61	0.00
	MHS h=12 88–07						
Model	I	II	III	IV	V	VI	VII
I	0.00	-1.86	-0.42	-0.67	-0.16	-0.64	-1.86
II	1.86	0.00	-0.11	-0.30	0.20	-0.32	-1.66
III	0.42	0.11	0.00	-1.04	1.54	-0.32	-1.02
IV	0.67	0.30	1.04	0.00	8.12	-0.08	-0.95
V	0.16	-0.20	-1.54	-8.12	0.00	-0.74	-1.35
VI	0.64	0.32	0.32	0.08	0.74	0.00	-0.96
VII	1.86	1.66	1.02	0.95	1.35	0.96	0.00

Note: Numbers are modified Diebold–Mariano (DM) test statistics which follow a t–distribution with $n-1$ degrees of freedom. Here, $n = 324$ is the number of out–of–sample forecasts. $H_0 : DM = 0$ can be rejected on the 5% level if the test statistic exceeds 1.97 in absolute values (two–sided test). A negative number means that the model in row i has a lower measured deviation than the model in column j. The first part captures results based on the whole sample, whereas the second and third part contain results for two sub–samples split in the end of 1987.

Table 5: Modified Diebold–Mariano test on deviations of *MHS* and forecasting models

Turning first to table 3, the left part gives results for the approximation of *SPF h=1*. Considering the whole sample, Model IV clearly shows negative values for the modified Diebold–Mariano test statistic throughout and, hence, dominates the other out–of–sample forecasts. Moreover, the approximation error is even significantly lower when compared to forecasts obtained by Models I and II. This means that, for *SPF h=1*, learning by signal–extraction clearly gives a better approximation of survey expectations than a simple backward–looking forecasting scheme which is given by Model I. Moreover, it also outperforms recursive least squares learning of coefficients, which is represented by Model II. Interestingly, it also performs much better than the recursively estimated models V and VI. Also note that all learning models III to VI yield a closer approximation of *SPF h=1* than the models which are not characterized by signal extraction. As argued in section 2.2, rational expectation formation is a poor proxy for survey expectations. Splitting the sample does not alter the results. Considering *SPF h=4*, which has a forecasting horizon of one year, it becomes apparent that Model III – the simplest signal extraction model – yields the best approximation. The difference here is even significant with the exception of Model I. This basically remains true for the first sub–sample. However, during the moderate period after 1987 the naive model proxies *SPF h=4* closest but if tested against Models III to VI the difference is not significant. Now turning to the left part of table 4, it is apparent that Model V yields the approximation closest to *LIV h=1*. Again, it outperforms Models I and II significantly when the test is based on the whole sample and the first sub–sample. Survey expectations from *LIV h=1* cannot be approximated by rational expectations which perform worst of all models. Looking at the right panel, results are mostly insignificant. For the whole sample period recursive least squares learning seems to yield the smallest deviation from *LIV h=2* and, again, rational expectations perform worst. During the period of disinflation, however, signal extraction gives the best description of expectation formation as Model IV performs best in the first sub–sample. The second sub–sample confirms the results found for the whole sample period. Coming now to table 5, which contains results for *MHS h=12*, findings are rather mixed. During the whole period, the recursively estimated Model VI gives the closest approximation of *MHS h=12*. Thus, one could conclude that, also in this case, signal extraction provides the best explanation for survey expectations. However, results are not significantly better than those obtained from rational expectations or naive and simple autoregressive forecasting schemes. Moreover, when taking a look at the first sample period, rational expectations seem to give the best approximation for *MHS h=12*. When looking at the second sub–sample, it is apparent that the naive forecasting scheme outperforms the other models.

To sum up, signal extraction gives a pretty good approximation of the expectation formation process. This is in particular the case for *SPF h=1* and *LIV* *h=1*. Here, learning by signal extraction generally outperforms other forecasting schemes. In particular, it leads to a significantly better approximation of expectation formation than recursive least squares learning. However, it remains unclear whether, expectations are even better characterized by signal extraction models whose estimated parameters are updated over time. Results for *MHS h=12* suggest that expectations of households – in opposition to experts – do not emerge from signal–extraction.

5 Heterogeneous Expectations

Having argued that signal–extraction gives the best approximation of expectation formation processes, it will now be important to show that these learning schemes indeed give a close and valuable explanation of survey expectations. One of the suggestions of sections 3.3 and 4.3 is the relevance of heterogeneity for an explanation of inflation expectations, as none of the models has so far been able to explain expectation formation perfectly. Consequently, each of the forecasting models of section 4 may play a role in aggregate expectation measures. However, we can estimate how important the respective model is for an explanation of survey expectations. Moreover, it will be possible to test if a weighted average of different model forecasts made in section 4.3 matches survey expectations arbitrarily closely. The weights are chosen such that the sum of squared deviations υ_t^2 of the linear model $\pi^e_{t+h|t} = \sum_{i=I}^{VII} \beta_i \pi^{f,i}_{t+h|t} + \upsilon_t$ is minimized with respect to β_i under the restrictions $\sum_{i=I}^{VII} \beta_i = 1$ and $0 \leq \beta_i \leq 1 \; \forall \; i$. The resulting estimates are presented in table 6. Additionally, the explanatory power of the respective linear model for survey expectations is provided by the R^2 and a Ljung–Box Q–test for autocorrelation is also given in the last two columns. All results are presented for the same sub–samples as before.

A first look at the weights for Models III to VI reveals that more than half of the participants seem to use a signal extraction type forecasting scheme. Interestingly, the results are quite robust across different surveys, although, they comprise quite different target variables and various forecasting horizons. The differences, however, occur between the two sub–samples of each survey which is in line with a time–varying behavior of respondents. The exception with respect to the estimation results is, again, *MHS* where learning plays no prominent role. Leaving *MHS* aside for the moment, it can be observed that the recursively estimated models V and VI attain zero weight in all cases except *LIV h=1*. Considering the two sub–samples separately, however, confirms the estimates of the other surveys. Results presented by Branch and Evans (2006) also point in this direction. One important result is given by the fact that signal extraction plays a prominent role in explaining survey

		I	II	III	IV	V	VI	VII	R^2	$Q(1)$	$Q(4)$
	80–07	0.00	0.00	**0.09**	**0.65**	0.00	0.00	**0.26**	0.86	0.00	0.00
SPF h=1	80–87	0.00	0.00	**0.35**	**0.47**	0.00	0.00	**0.19**	0.88	0.86	0.12
	88–07	0.00	0.00	0.00	**0.71**	0.00	0.00	**0.29**	0.34	0.00	0.00
	80–07	**0.40**	0.00	**0.59**	0.00	0.00	0.00	**0.01**	0.87	0.00	0.00
SPF h=4	80–87	**0.18**	0.00	**0.82**	0.00	0.00	0.00	0.00	0.86	0.03	0.21
	88–07	**0.81**	0.00	**0.11**	**0.07**	0.00	0.00	**0.02**	0.43	0.00	0.00
	80–07	0.00	**0.39**	0.00	0.00	0.00	**0.27**	**0.34**	0.68	0.00	0.00
LIV h=1	80–87	0.00	**0.37**	0.00	**0.52**	0.00	0.00	**0.11**	0.46	0.73	0.86
	88–07	**0.52**	0.00	0.00	0.00	0.00	**0.02**	**0.46**	0.42	0.00	0.01
	80–07	0.00	**0.46**	**0.34**	0.00	0.00	0.00	**0.21**	0.79	0.00	0.00
LIV h=2	80–87	0.00	**0.47**	**0.38**	**0.07**	0.00	0.00	**0.08**	0.65	0.70	0.21
	88–07	**0.57**	0.00	0.00	0.00	0.00	0.00	**0.43**	0.56	0.00	0.00
	80–07	**0.19**	0.00	0.00	0.00	0.00	**0.33**	**0.48**	0.79	0.00	0.00
MHS h=12	80–87	**0.20**	0.00	0.00	0.00	0.00	**0.20**	**0.60**	0.85	0.00	0.00
	88–07	**0.13**	0.00	**0.28**	0.00	0.00	**0.28**	**0.30**	0.16	0.00	0.00

Note: Numbers represent weights of the respective forecast in the survey forecast. These weights are chosen such that the sum of squared deviations υ_t of the linear model $\pi^e_{t+h|t} = \sum_{i=I}^{VII} \beta_i \pi^{f,i}_{t+h|t} + \upsilon_t$ is minimized under the restrictions $\sum_{i=I}^{VII} \beta_i = 1$ and $0 \leq \beta_i \leq 1 \forall i$. The columns labeled $Q(1)$ and $Q(4)$ contain p–values for a Ljung–Box Q–test for autocorrelation up to 1 and 4 periods, respectively. The way it is conducted here, the test does not account for additional uncertainty contained in the out–of–sample forecasts $\pi^{f,i}_{t+h|t}$ which enter the model.

Table 6: Estimated weights

expectations especially during the Volcker period. For *SPF*, signal extraction makes up for over 80% during the first sub–sample, whereas during the second period it attains a weight of about 70% for *SPF h=1* and only 18% for *SPF h=4*. In the case of *LIV* this tendency is even more pronounced, as fractions change from more than one half for the early period to virtually zero percent for the period after 1987. Furthermore, rational expectations get a weight which is below 50% for all surveys except *MHS*, which is in line with the results from above. Interestingly, the fraction of rational respondents is always higher for the second sub–sample. At the same time, however, also the share of naive forecasters increases for all surveys except *SPF h=1*. Note, that in this case simple backward–looking behavior represented by Models I and II does not contribute to survey expectations. Now turning to *MHS*, results suggest that there is no prominent role for learning behavior. Agents seem to be either naive forecasters or rational. Only one third is found to use Model VI for forecasting.

Looking at the last three columns of table 6, it becomes apparent that the explanatory power of the estimated linear relationships is quite high. The R^2 suggests that more than 80% of the variation in case of *SPF* and about 70% of the variation in case of *LIV* can be explained. Taking sub–samples into account yields another interesting result. The explanatory power of the estimated relationships is higher for the early disinflation period compared to the years after 1987. Moreover, there are no signs of autocorrelation in the estimations covering the Volcker period[12]. Although R^2 is also fairly high for *MHS* during the first period, the residual is still autocorrelated. Dynamics of the second period are less precisely described by heterogeneous expectations. This finding basically confirms the results presented in tables 2 and 5.

On the whole, the concept of heterogeneous expectations with a prominent role for signal extraction is well suited to explain survey measures of inflation expectations. During phases of disinflation, the model has more explanatory power than in tranquil periods, as, during the Volcker period, the R^2 is higher, the unexplained part is free of autocorrelation and the role of signal extraction is even more prominent.

6 Conclusions

In a first step, I have shown that the behavior of surveys on inflation expectations is not compatible with the concept of rational expectations. Survey expectations are characterized by temporary bias and considerable persistence of forecast errors. Many theoretical studies emphasize the importance of persistence of inflation expectations for the dynamics of the inflation rate. Moreover, theoretical models that assume rational expectations unrealistically predict a jump of inflation expectations following a change of the inflation target. Most importantly, such a behavior of inflation expectations cannot explain why disinflation is costly in a purely forward–looking framework. A possible solution is given by the signal extraction framework presented in this paper. Here, it is important to note that agents are not assumed to be completely ignorant. By contrast, they are confronted with a difficult forecasting problem. The reason is that the inflation target pursued by the central bank is not directly observable but has to be estimated from a noisy signal. The solution to this signal extraction problem is given by the Kalman filtering framework which constitutes the learning rule of private agents. It can be shown that it is possible to fit such a model to inflation expectations of *SPF* and *LIV*. The in–sample results suggest rather slow learning of trends which can explain the sluggishness of U.S. inflation expectations.

In a next step, I conduct an out–of–sample forecasting exercise to simulate a forecaster that solves the signal extraction problem by Kalman filtering. Different

[12]The way it is conducted here, the test does not account for additional uncertainty contained in the out–of–sample forecasts $\pi^{f,i}_{t+h|t}$ which enter the model as explanatory variables. Therefore, standard parameter distributions and test statistics do not apply in this case. However, I use standard autocorrelation tests to test for systematic behavior of υ_t. This can be justified by the fact that the test is constructed with a null hypothesis of no autocorrelation and, thus, will reject too often if additional estimation uncertainty is not taken account of.

types of learning by Kalman filtering are contrasted with a naive forecaster, learning by recursive least squares and a rational forecaster. It turns out that learning by Kalman filtering approximates U.S. survey expectations closest – at least during the presidency of Volcker. This holds true for several surveys comprising several target variables. Finally, in the spirit of heterogeneous expectations, I construct a weighted average of the employed forecasting schemes. It turns out that the concept of heterogeneous expectations with a prominent role for signal extraction is well suited to explain survey measures of inflation expectations during the Volcker period. The R^2 is high, the unexplained part is free of autocorrelation and the role of signal extraction is even more prominent. On the whole, learning in an uncertain environment provides a good explanation for the sluggishness of inflation expectations. Moreover, a large fraction of agents seems to solve some signal extraction problem during phases of disinflation. Consequently, in a general equilibrium framework, it will be worthwile to replace the common notion of rational expectations and to model the expectation formation process in a way consistent with signal extraction behavior.

Appendix

A Reparameterization of Variables in Section 3.3.2

The parameters contained in ψ are reparameterized such that they obey the theoretical restrictions. The parameter vector estimated by maximum likelihood is denoted by θ. It is transformed by a vector of functions $\mathbf{g}(\theta)$ in the following way:

$$\psi = \begin{pmatrix} \lambda \\ \rho \\ \sigma_\varepsilon^2 \end{pmatrix} \equiv \mathbf{g}(\theta) = \begin{pmatrix} g_1(\theta) \\ g_2(\theta) \\ g_3(\theta) \end{pmatrix} = \begin{pmatrix} \exp(\theta_1) \\ \Phi(\theta_2) \\ \exp(2\theta_3) \end{pmatrix}. \tag{9}$$

In the second step we need to calculate the standard errors of the estimates. It can be shown that the transformed estimates are asymptotically normal with estimated variance $\widehat{\text{var}}(\widehat{\psi}) = G_{\widehat{\theta}}\, \widehat{\text{var}}(\widehat{\theta})\, G'_{\widehat{\theta}}$.[13] This yields the following adjustment matrix of first derivatives $G_{\widehat{\theta}}$[14].

[13]See for example Kim and Nelson (1999), chapter 2.

[14]The calculation of the standard error of the transformed estimates was done with the *Delta Method* which relies on first–order Taylor expansions of non–linear functions. For an overview compare Davidson and MacKinnon (2004) chapter 5.6.

$$
\begin{aligned}
G_{\widehat{\theta}} &= \begin{pmatrix} \frac{\partial g_1(\widehat{\theta})}{\partial \theta_1} & \cdots & \frac{\partial g_1(\widehat{\theta})}{\partial \theta_3} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_3(\widehat{\theta})}{\partial \theta_1} & \cdots & \frac{\partial g_3(\widehat{\theta})}{\partial \theta_3} \end{pmatrix} \\
&= \begin{pmatrix} \exp(\widehat{\theta}_1) & 0 & 0 \\ 0 & \phi(\widehat{\theta}_2) & 0 \\ 0 & 0 & 2\exp(2\widehat{\theta}_3) \end{pmatrix}.
\end{aligned}
$$

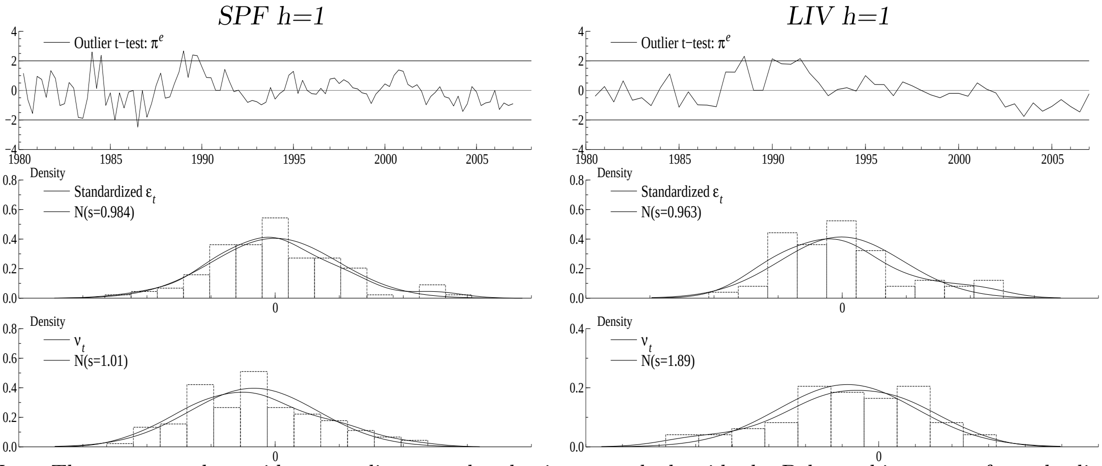
B Reparameterization of Variables in Section 4

The parameters contained in ψ are reparameterized such that they obey the theoretical restrictions. The parameter vector estimated by maximum likelihood is denoted by θ. It is transformed by a vector of functions $\mathbf{g}(\theta)$ in the following way:

$$
\psi = \begin{pmatrix} \sigma^2_{\eta,2} \\ \sigma^2_\kappa \\ \lambda \\ \rho \\ \sigma^2_\varepsilon \end{pmatrix} \equiv \mathbf{g}(\theta) = \begin{pmatrix} g_1(\theta) \\ g_2(\theta) \\ g_3(\theta) \\ g_4(\theta) \\ g_5(\theta) \end{pmatrix} = \begin{pmatrix} \exp(2\theta_1) \\ \exp(2\theta_2)(1-\Phi(\theta_4)^2) \\ \exp(\theta_3) \\ \Phi(\theta_4) \\ \exp(2\theta_5) \end{pmatrix}. \tag{10}
$$

C Diagnostics of the Learning Model

Figure 5 shows some diagnostics for the learning model. It depicts the respective standardized irregular component along with the associated histogram. The respective third panel shows a histogram of forecast errors for comparison.



Note: The upper graph provides an outlier t–test by plotting smoothed residuals. Below, a histogram of standardized errors ε_t can be found. The last panel of the respective graph plots the distribution of observed forecast errors ν_t.

Figure 5: Diagnostics learning model

D Recursive Parameter Estimates

Figures 6 to 10 depict structural parameter estimates (left panel) along with steady–state gain parameters taken from the state vector of the system described by equations (2) to (4) (right panel). The upper part of the respective graph shows estimates from Model V whereas estimates for Model VI are presented in the lower panel.

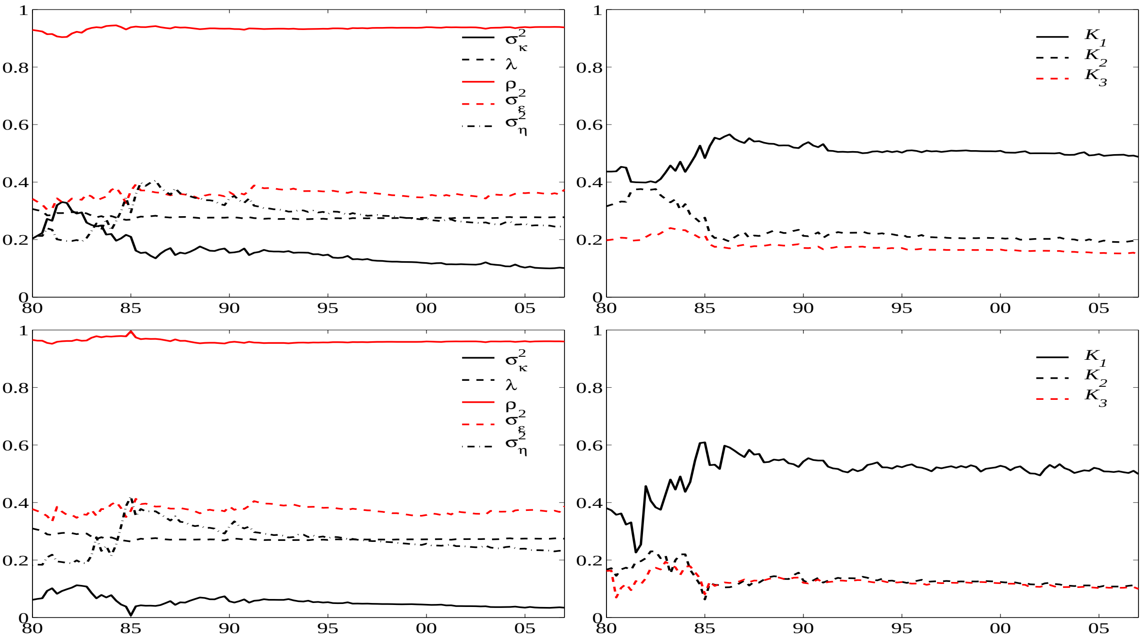


Figure 6: Recursively estimated parameters, *annualized quarterly GDP inflation h=1*

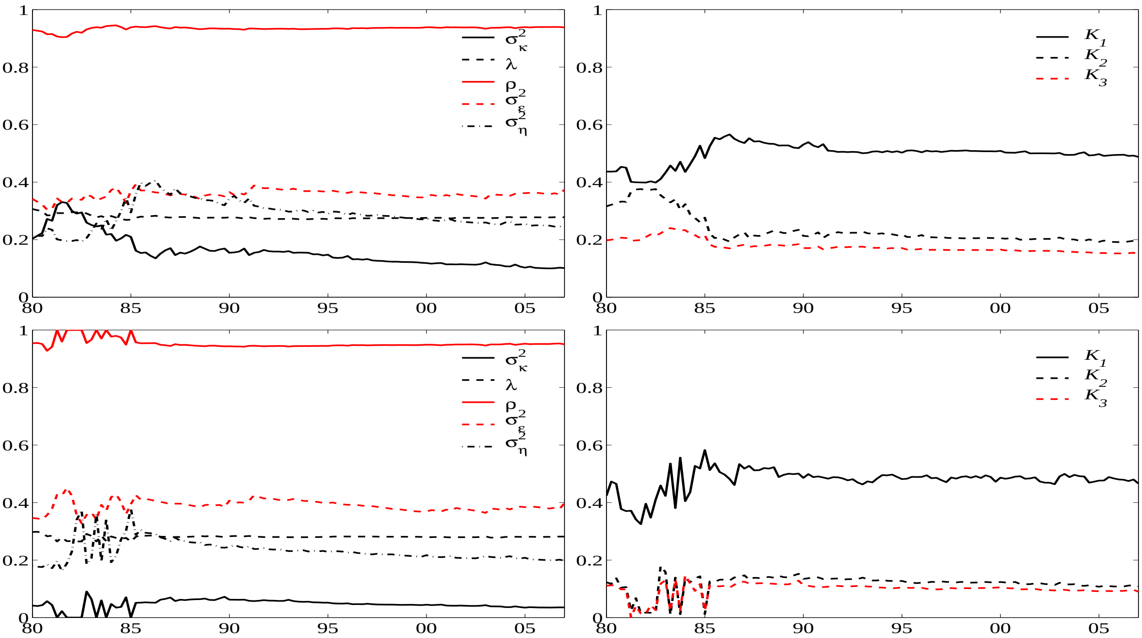


Figure 7: Recursively estimated parameters, *average annualized 4 quarter GDP inflation h=4*

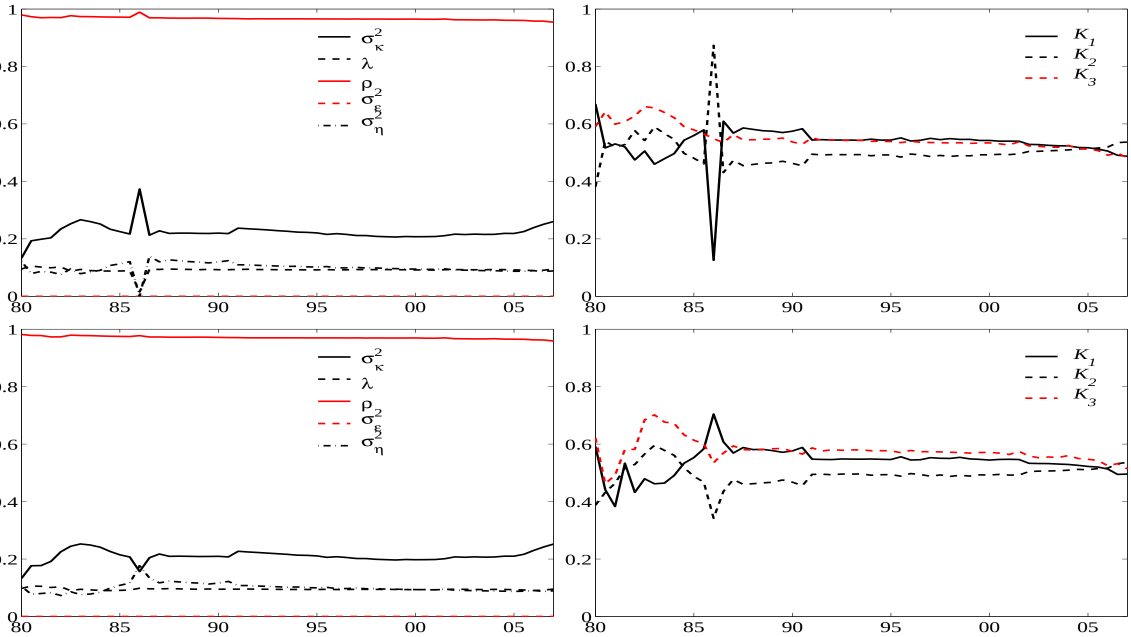


Figure 8: Recursively estimated parameters, *annualized 6 month CPI inflation h=1*

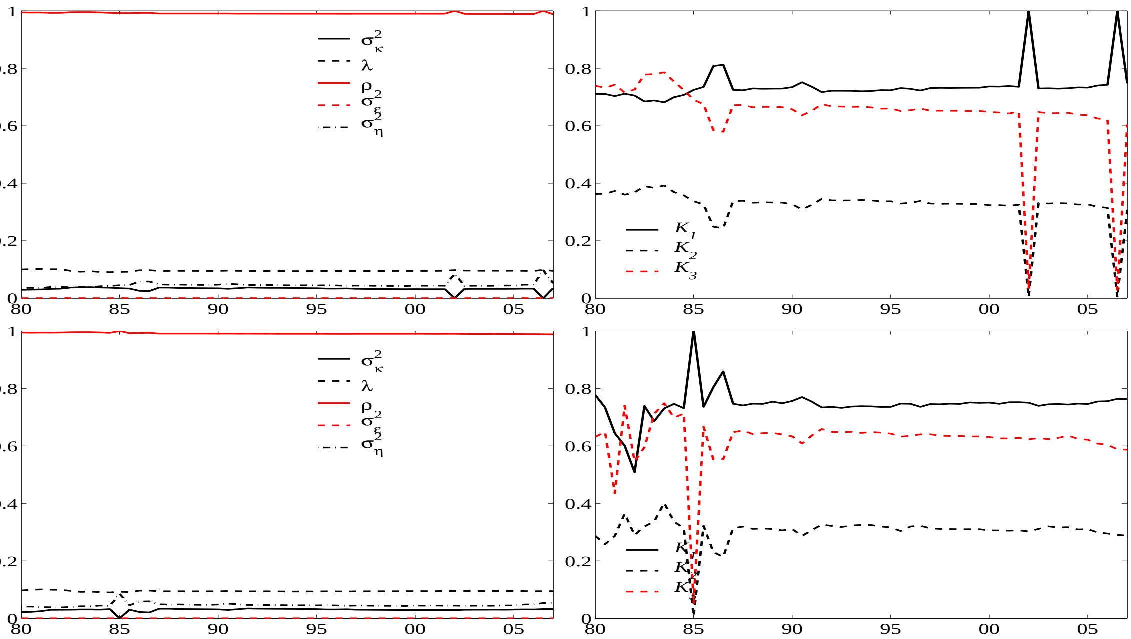


Figure 9: Recursively estimated parameters, *12 month CPI inflation h=2*

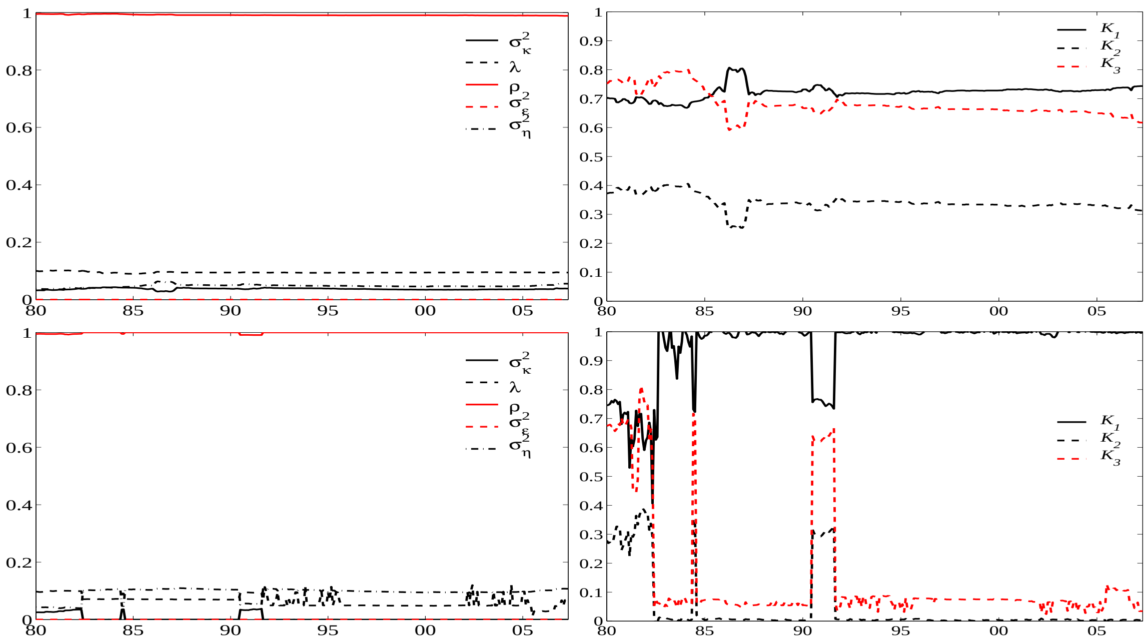


Figure 10: Recursively estimated parameters, *12 month average CPI inflation h=12*

References

Andolfatto, D., and P. Gomme (2003): "Monetary policy regimes and beliefs," *International Economic Review*, 44(1), 1–30.

Andolfatto, D., S. Hendry, and K. Moran (2002): "Inflation expectations and learning about monetary policy," *Bank of Canada Working Paper*, 30.

——— (2006): "Are inflation expectations rational?," manuscript.

Branch, W. A. (2004): "The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations," *The Economic Journal*, (114), 592–621.

Branch, W. A., and G. W. Evans (2006): "A simple recursive forecasting model," *Economics Letters*, 91, 158–166.

Cogley, T., and A. M. Sbordone (2006): "Trend inflation and inflation persistence in the New Keynesian Phillips curve," *Federal Reserve Bank of New York Staff Reports*, (270).

Cukierman, A., and A. H. Meltzer (1986): "A theory of ambiguity, credibility, and inflation under discretion and asymmetric information," *Econometrica*, 54(5), 1099–1128.

Cumby, R. E., and J. Huizinga (1992): "Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions," *Econometrica*, 60(1), 185–195.

Davidson, R., and J. G. MacKinnon (2004): *Econometric theory and methods*. Oxford University Press.

Diebold, F. X., and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Dossche, M., and G. Everaert (2005): "Measuring inflation persistence – A structural time series approach," ECB Working Paper No. 495.

Erceg, C. J., and A. T. Levin (2003): "Imperfect credibility and inflation persistence," *Journal of Monetary Economics*, 50, 915–944.

Evans, G. W., and S. Honkapohja (2001): *Learning and expectations in macroeconomics*. Princeton University Press.

——— (2003): "Adaptive learning and monetary policy design," *Journal of Money, Credit and Banking*, 35(6), 1045–1072.

Evans, M., and P. Wachtel (1992): “Inflation regimes and the sources of inflation uncertainty,” *Journal of Money, Credit and Banking*, 25(3), 475–511.

Hamilton, J. D. (1994): *Time series analysis.* Princeton University Press.

Harvey, A. C. (1989): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Harvey, A. C., and S. J. Koopman (1992): “Diagnostic checking of unobserved components time series models,” *Journal of Business & Economic Statistics*, 10(4), 377–389.

Harvey, D. I., S. J. Leybourne, and P. Newbold (1997): “Testing the equality of prediction mean squared errors,” *International Journal of Forecasting*, 13, 281–291.

Kim, C.-J., and C. R. Nelson (1999): *State–space models with regime switching.* MIT Press, Cambridge.

Kozicki, S., and P. A. Tinsley (2005): “Permanent and transitory policy shocks in an empirical macro model with asymmetric information,” *Journal of Economic Dynamics & Control*, 29(11), 1985–2015.

Muth, J. F. (1961): “Rational expectations and the theory of price movements,” *Econometrica*, 29(3), 315–335.

Nunes, R. C. (2004): “Learning the inflation target,” manuscript, Universitat Pompeu Fabra.

Pajfar, D., and E. Santoro (2007): “Heterogeneity and learning in inflation expectation formation: An empirical assessment,” manuscript, University of Cambridge.

Roberts, J. M. (1997): “Is inflation sticky?,” *Journal of Monetary Economics*, 39, 173–196.

Weber, A. (2007): “Heterogeneous expectations, learning and European inflation dynamics,” Deutsche Bundesbank Discussion Paper No. 16/07.