Heterogeneity, Learning and Information Stickiness in Inflation Expectations*

Damjan Pfajfar[†]
CentER, EBC, University of Tilburg

Emiliano Santoro[‡]
University of Copenhagen
and University of Trento

June 2, 2009

Abstract. In this paper we propose novel techniques for the empirical analysis of adaptive learning and sticky information in inflation expectations. These methodologies are applied to the distribution of households' inflation expectations collected by the University of Michigan Survey Research Center. In order to account for the evolution of the cross-section of inflation forecasts over time and to measure the degree of heterogeneity in private agents' forecasts, we explore time series of percentiles from the empirical distribution. Our results show that heterogeneity is pervasive in the process of inflation expectation formation. We identify three regions of the distribution that correspond to different underlying mechanisms of expectation formation: a static or highly autoregressive region on the left hand side of the median, a nearly rational region around the median and a fraction of forecasts on the right hand side of the median formed in accordance with adaptive learning and sticky information.

JEL: E31, C53, D80

Keywords: Heterogeneous Expectations, Adaptive Learning, Sticky Information, Survey Expectations

*We would like to thank Sean Holly, Seppo Honkapohja and Hashem Pesaran for guidance and helpful discussion. We would also like to thank Steffan Ball, Jagjit Chada, Daniele Massacci, Linda Rousova and seminar participants at Keele University, Cambridge University, Catholic University in Milan, WEHIA 2006, St. Louis FED Learning Week 2006, EDGE Jamboree 2006, MMF 2006, EEA Annual Meeting 2007, and 2nd Ifo Conference on Survey data for their comments. All remaining errors are ours. The paper was previously circulated under the title "Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment".

[†]Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands. *E-mail*: D.Pfajfar@uvt.nl. *Web*: http://center.uvt.nl/staff/pfajfar/.

[‡]University of Copenhagen, Department of Economics, Studiestræde 6, 1455 Copenhagen K, Denmark *E-mail*: emiliano.santoro@econ.ku.dk.

Introduction

This paper deals with the analysis of private agents' inflation expectations. Despite the fact central banks claim that managing inflation expectations is one of the most important prerequisites for attaining price stability and promoting sustainable growth, still very little is known about consumers' process of expectation formation. As noted by Bernanke (2007), reported private sector inflation expectations are important because they signal future inflationary risks and provide indications about agents' perception of these risks. Private inflation expectations often diverge from those of the central bank, and may represent a distinct source of information as well as a potential intermediate target for the conduct of monetary policy. We argue that valuable information can be extracted by analyzing the distribution of households' inflation forecasts, whereas a substantial number of studies have typically focused on measures of central tendency, such as the mean or the median forecast.

Our analysis is centered on the development of novel techniques for the assessment of different mechanisms of expectation formation which have been recently advanced in the theoretical literature. The common trait of these theories is to relax the hypothesis of perfectly informed agents, as assumed in the rational expectations paradigm. Some of these theories postulate the existence of informational frictions generating sticky expectations, while others conjecture that agents might act as econometricians when forecasting. The latter approach, widely known as adaptive learning, is extensively discussed in Evans and Honkapohja (2001). As to sticky expectations, a number of papers (e.g. Carroll 2003a, and Mankiw and Reis, 2002) show how to generate time dependent rules under which expectations are updated only at fixed intervals. Carroll (2003a, 2003b) proposes an epidemiological framework where consumers update their inflation expectations from the media, which are assumed to transmit the expectations of the professional forecasters. Mankiw and Reis (2002) suggests that agents update information more frequently when inflation matters. We put forward novel techniques for the empirical assessment of adaptive learning and inattentiveness in inflation expectations.

We apply our methodologies, along with traditional tests for rational and adaptive expectations,[1] to the distribution of households' inflation expectations collected by the University of Michigan Survey Research Center. Our focus on the cross section of private agents' forecasts is aimed at assessing different sources of heterogeneity in the process of expectation formation. In order to account for the evolution of the cross-section of inflation forecasts over time, we compute percentiles of the empirical distribution in each period. Therefore, we retrieve monthly time series for each percentile, which convey information on the individuals comprised in different parts of the empirical density. We find that the null hypothesis of rationality cannot be rejected just for few percentiles, which are generally placed around or slightly above the median forecast. Only less than 10% of households' forecasts reflect regular information updating. We augment the epidemiological framework proposed by Carroll (2003a, 2003b) to account for the impact of the level of inflation on the frequency of information updating. The resulting framework is based on the assumption that agents are more likely to regularly update their information set

[1] See, for a review of these tests, Pesaran (1987), Mankiw, Reis, and Wolfers (2004), and Bakhshi and Yates (1998).

when inflation matters. This is found to be a plausible explanation for the forecast range in the upper end of the distribution, where greater attentiveness is paid in periods of high and volatile inflation.

We put forward a novel technique to detect the presence of adaptive learning in the distribution of forecasts. The initialization of the learning algorithm is of crucial importance in the estimation of the gain parameter that indexes the speed of learning. Previous estimation techniques of models under adaptive learning (Milani, 2007) have generally been pursued by splitting the time series into two subsamples. Thus, the first subsample is used to estimate the set of initial values in the Perceived Law of Motion (PLM). In turn, the initial values represent the starting point for the recursive estimation of the gain parameter in the second subsample. Clearly, the main practical inconvenience of this approach is that it does not allow the researcher to fully exploit the data available. In addition, this approach still bears the risk that learning dynamics could just result as a statistical artifact due to non-optimal initialization. Our approach abstracts from this criticism, as we search for the combination of initial values and gain parameter that provides the closest explanation of the empirical density, thus preserving the sample structure and optimizing the initialization procedure.[2] Our results suggest that consumers' forecasts on the right hand side of the median (RHS) display adaptive learning, whereas forecasts on the left hand side of the median (LHS) do not exhibit such behavior.

We propose an alternative mechanism of expectation formation, whereby households are assumed to update their forecasts with respect to (expected) future errors, which are reflected in the difference between their forecasts and the predictions of the professional forecasters. This mechanism draws on the epidemiological view advanced by Carroll (2003a), and represents a combination of adaptive learning and sticky information.

Additional time series techniques take into account a wider set of explanatory variables for inflation forecasts and confirm a significant degree of heterogeneity and asymmetry in the underlying information structure. The forecast range at the center of the distribution is generally unbiased. However, our results suggest that forecasts on the LHS are highly static and entail systematic errors. It can be argued that expectations in this region of the empirical density are stable around certain digits and that they do not reflect movements in any of the macroeconomic variables considered as relevant for the forecasting process. Conversely, RHS forecasts tend to overreact to changes in actual inflation. These findings are in line with the evidence carried out by Curtin (2005), who points out that negative changes in inflation exert twice the impact as positive changes. As noted above, RHS inflation expectations are consistent with adaptive behavior. In this forecast range information is updated only from time to time.

Three different roots of heterogeneity have been traditionally explored in the literature. Heterogeneous forecasts might be the consequence of agents (i) employing different models,[3] (ii) relying on different information sets or (iii) entailing different capacities to process in-

[2] Orphanides and Williams (2003, 2005a, 2005b) and Milani (2007) have provided some empirical support for adaptive learning dynamics in DSGE models. They provide an estimate of the gain parameter.

[3] Namely, agents could have different underlying assumptions about the structure of the economy or different parameterizations (or priors) of the same model.

formation. Some theoretical studies have introduced heterogeneous expectations in standard macroeconomic models, such as in the New Keynesian model (Branch and McGough, 2006). Branch (2004, 2007) assesses the importance of the first two roots of heterogeneity and finds that data are consistent with both of them. He replicates some of the inherent characteristics of the Michigan Survey Households' Expectations empirical distribution, designing a switching mechanism between alternative models of prediction and different frequencies of information updating, based on their relative historical performance. Nevertheless, as mentioned above, most of these studies only focus on measures of central tendency to assess the degree of heterogeneity in private forecasts. We show that this approach entails some fundamental fallacies if forecast distributions are not time invariant and display substantial asymmetry. We provide evidence on the cross-sectional dynamics of inflation forecasts, showing that different regions of the distribution reflect different forecasting mechanisms. Compared to Branch (2004, 2007), we allow for a wider range of forecasting mechanisms, including a combination of adaptive learning and information stickiness.

The remainder of the paper reads as follows: Section 1 overviews the Survey of Consumer Attitudes and Behavior; Section 2 reports some preliminary descriptive statistics; Section 3 focuses on the methodology developed in the paper, reporting some applications of our techniques on adaptive learning and informational stickiness; Section 4 summarizes and discusses the main empirical results; last section concludes.

1. The Survey of Consumer Attitudes and Behavior

The Survey of Consumer Attitudes and Behavior, conducted by the Survey Research Center (SRC) at the University of Michigan, has been available on a monthly basis since January 1978. The survey comprises a cross-section of about 500 households per month.[4] After the first interview, each respondent is re-interviewed within six months. The sampling method is designed in a way that, in any given month, approximately 45% of prior respondents are re-interviewed, while the remaining 55% are new households. There are two relevant questions about price level changes: (i) firstly, households are asked whether they expect prices to go up, down or to stay the same in the next 12 months; (ii) secondly, they are asked to provide a quantitative answer about the expected change.[5]

Publicly available data are summarized in intervals (e.g. "go down", "stay the same or down", go up by $1-2\%$, $3-4\%$, 5%, $6-9\%$, $10-14\%$, $15+\%$). There might be some confusion about the category "stay the same or down". We follow Curtin (1996) and regard this response as 0. A word of caution is in order for households that expect prices to go up without providing any quantitative statement. In this case, we redistribute their response across the six ranges of price change, depending on their relative size with respect to the overall sample size. Only a negligible proportion of "do not know" responses is excluded.

[4] A peak of 1,479 households occurs in November 1978 and a minimum of 492 in November 1992. An average of approximately 500 respondents is sampled since January 1987.

[5] In case the respondent expects prices to stay the same, the interviewer must make sure that the respondent does not actually have in mind a change in the price level which is assimilable to the one measured at the time of the interview.

As agents report unbounded inflation forecasts, we need to determine point at both ends of the distribution beyond which observations should be excluded.[6] Curtin (1996) suggests two alternative truncations, namely at -10% and $+50\%$ and at -5% and $+30\%$.[7] The analysis carried out on the distribution obtained from different truncation intervals does not produce any major discrepancy. Thus, in the remainder of the paper we only present evidence derived under the second truncation rule.

2. A Preliminary Look at the Data

We consider the time window between 1978.01 and 2005.02. Within this period, we explore the dynamic pattern of the moments of the Michigan Survey Households' Expectations distribution (MSHE hereafter). To account for the presence of a structural break, we pursue a parallel investigation on two subsamples, namely pre- and post-1988.12. This choice allows us to take into adequate account the highly inflationary period characterizing the first part of the sample and the subsequent disinflation. No significant differences can be highlighted.

2.1. Descriptive Statistics. In the remainder time series on expectational variables are plotted at the realized date and not at the time the forecast has been produced. Figure 1(a) plots mean and median of the MSHE distribution against actual inflation. It is evident how both measures of central tendency constantly underestimate the rise in inflation in the first part of the sample, although the forecasting performance improves remarkably during the subsequent disinflation. This is probably due to the fact that the Federal Reserve (FED) has acquired more credibility in fighting inflationary pressures. In the post-1988 subsample, expectations appear to be quite stable, although they almost systematically over-estimate inflation. Furthermore, we can observe how expectations fail to account for the marked rise in price level during the first Gulf War, by reacting with a consistent delay. This over-reaction is also present after 9/11, although with opposite sign.

Higher empirical moments, together with the median forecast, are reported in Figures 1(b)-(c). Not surprisingly, higher expectations are usually associated with higher volatility. Opposite evidence holds for skewness and kurtosis, although both statistics tend to fluctuate around a lower level in the highly inflationary period (opposite evidence holds for the variance).

Figure 1(d) displays the 25^{th}, the 50^{th} and the 75^{th} percentiles. Marked differences can be observed in the degree of volatility of different parts of the distribution. The 75^{th} percentile appears to be remarkably stable after 1988, compared to the other series. At the same time, the median forecast reacts less and with a marked delay to the inflationary pressures triggered by the first Gulf War, while it is rather reactive in the aftermath of the 9/11. A possible explanation for agents in the center of the distribution to react more is that they perceived a threat of deflation as credible.

[6] It is important to recall that the exact specification of the truncation rule only influences the mean and the variance of the distribution, but has no effect on the median. It is also important to take into account that the upper tail of the distribution is not only long but also sparse, frequently with large gaps between observations. Technical considerations regarding the cut-off procedure are outlined in Curtin (1996).

[7] Curtin (1996) also suggests that there is not compelling evidence supporting the choice of one truncation rule over the other.

Insert Figure 1 about here

Figure 1(e) reports the mean of the MSHE distribution and of the Survey of Professional Forecasters (SPF) forecasts against actual inflation. It is striking how professional forecasts, generally more accurate in the second part of the sample, are more biased than households' expectations during the period of high inflation. The two predictions are remarkably similar from 1984 to 1990 and from this point onward the SPF provides a more accurate prediction.

Figures 1(f)-(g) report higher moments of the distribution against two different cycle indicators.[8] The variance displays a marked counter-cyclical behavior, while skewness and kurtosis are pro-cyclical. Moreover, the third and fourth moments display higher variability in the post-1988 period. Kurtosis exhibits increasing variability when the cycle peaks. This reflects rising uncertainty about the future after the turning point. The skewness dynamics is qualitatively similar to the one featuring kurtosis. This evidence signals the existence of a long and sparse positive tail characterized by high variability.[9]

3. Percentile Time Series Analysis

The remainder of the paper develops formal procedures to assess recent theories of expectation formation, such as information stickiness and adaptive learning. Particular attention is paid to the information retrievable from the entire distribution of responses. The approach pursued, aimed at tracking the evolution of the cross-sectional dimension, relies on the use of time series of percentiles. Percentile time series analysis is not only motivated by its suitability to account for asymmetric responses in the MSHE distribution, but also by a more practical consideration. The panel under scrutiny is highly unbalanced, as every respondent is interviewed only twice. Computing percentiles for each year provides us with time series of equidistant statistics that describe the distribution both under a dynamic and a longitudinal perspective. We regard the expected change in price level in the next 12 months as a random variable $\left(\pi_{t|t+12}\right)$ with distribution $F(\cdot)$. The k^{th} percentile $\left(\pi^k_{t|t+12}\right)$ is the value below which $k\%$ of the responses lie.[10] Therefore, we retrieve a set of ordered statistics for each month, i.e. 99 time series of percentiles.[11]

A number of studies in the past have employed the mean or median forecasts from the Michigan Survey.[12] Implicitly, one motivation for focusing on measures of central tendency is to remove any idiosyncratic component in the cross-section of forecasts. This principle applies

[8] These consist of a HP detrended industrial production index (IPI) and an interpolated estimate of Kuttner (1994) model of multivariate Kalman filtering.

[9] Time series analysis on higher moments of the distribution of inflation forecasts confirms our visual impression.

[10] Thus $F(\pi^k_{t|t+12}) = k$.

[11] We are aware of the methodological limits implicit in this approach, as the survey is not conducted on the same households' throughout the time window considered. Nevertheless, some empirical (e.g. Pfajfar and Santoro, 2008a and Curtin, 2005) support the view that agents with analogous characteristics tend to behave similarly. Inflation forecasting is common in every-day life and not just when households are asked to provide their forecasts. We can argue that when one respondent is replaced by another with similar intrinsic characteristics, her information set is likely to be nested within the newcomer's one. This argument is in line with the conceptual structure of overlapping generation models.

[12] Often, the median forecast is preferred over the mean, given that extreme observations are not considered to be particularly informative. This is detailed in Mankiw, Reis, and Wolfers (2004).

when dealing with symmetric and unimodal densities. However, it can be shown that the median inflation forecast of the Michigan Survey may not be an appropriate measure of central tendency, given that both pooled and time series data display substantial asymmetry and multimodality.

3.1. Rationality Tests. The rational expectations hypothesis (REH) can be interestingly tested with survey expectations data[13] to allow for different degrees of forecast efficiency across the distribution of responses. To satisfy the REH, the forecasting procedure should not yield predictable errors. A test of bias can be applied by regressing the expectation error of each percentile on a constant.[14] This allows us to verify whether inflation expectations are centred around the right value:

$$\pi_t - \pi^k_{t|t-12} = \alpha + \varepsilon_t, \tag{1}$$

where π_t is inflation at time t and $\pi^k_{t|t-12}$ is the k^{th} percentile from the MSHE. The following regression represents a second test for rationality:

$$\pi_t = a + b\pi^k_{t|t-12} + \varepsilon_t, \tag{2}$$

where rationality implies that conditions $a = 0$ and $b = 1$ are jointly satisfied. Equation (2) can be simply augmented to test whether available information is fully exploited:

$$\pi_t - \pi^k_{t|t-12} = a + (b-1)\,\pi^k_{t|t-12} + \varepsilon_t. \tag{3}$$

Under the null of rationality, these regressions are meant to have no predictive power.[15]

Results. Regressions based on equation (1) suggest that only the $51^{st} - 55^{th}$ ($52^{nd} - 54^{th}$) percentile range is not biased at a 5% (1%) level of significance. Testing for bias has been often conducted on survey data. Among others, Croushore (1998), Roberts (1997), and Mankiw, Reis, and Wolfers (2004) test for rationality in the Michigan Survey. They focus on the mean and median forecast and tend to reject the null hypothesis.[16] We cannot reject the null of rationality for some percentiles placed slightly above the median. When splitting the sample into pre-1988 and post-1988, we find that forecasts between the 55^{th} and 63^{rd} percentile ($56^{th} - 62^{nd}$) are not biased at a 5% (1%) level of significance. In the $1989 - 2005$ period, forecasts in the $47^{th} - 50^{th}$ ($48^{th} - 50^{th}$) percentile range are not biased at a 5% (1%) level of significance. A greater share of rational forecasts can be estimated in the first subsample, when inflation is higher and induces a higher level of attentiveness. The second test for rationality (3) always leads to rejection of the null hypothesis of rationality.

[13] See Pesaran (1987), Mankiw, Reis, and Wolfers (2004) and Bakhshi and Yates (1998) for a review of these tests.

[14] See, for an application, Jonung and Laidler (1988) and Mankiw, Reis, and Wolfers (2004).

[15] An alternative test for rationality takes into account that inflation and inflation expectation data are I(1). The REH suggests that these series cointegrate, i.e. expectations errors are stationary. Moreover, the cointegrating vector has no constant terms and the coefficients on expected and actual inflation should be equal in absolute value (Bakhshi and Yates, 1998).

[16] Roberts (1997) tests for the REH in survey expectations. He concludes that both Michigan and Livingston forecasts display an intermediate degrre of rationality, being nor fully rational, neither entirely adaptive.

3.2. A further Investigation on the Forecast Error. The results reported above suggest that a substantial part of households' inflation forecasts are not rational. To explore further the nature and the determinants of the forecast error we estimate model (4). Regressing forecast errors on different determinants, such as changes in actual inflation and errors of the professional forecasters,[17] provides us with valuable information on the degree of heterogeneity in the data:

$$\begin{aligned} \pi_t - \pi^k_{t|t-12} &= \alpha + \beta\left(\pi_{t-13} - \pi^k_{t-13|t-25}\right) + \delta(\pi_t - \pi^F_{t|t-12}) + \boldsymbol{\gamma}\boldsymbol{\Delta}\mathbf{X}_t + \varepsilon_t, \\ k &= 1, ..., 99; \qquad \mathbf{X}_t = \begin{bmatrix} y_t & \pi_t & (i_t - r_t) \end{bmatrix}'. \end{aligned} \tag{4}$$

where $\pi^k_{t|t-12}$ denotes the the k^{th} percentile of the 12 months ahead expected change in prices, while $\pi^F_{t|t-12}$ denotes the mean of the 12 months ahead expected change in prices of the SPF. Moreover, y_t, π_t, i_t and r_t denote output gap, actual inflation, real short term interest rate (3-month t-bill coupon rate) and long term interest rate (10-year t-bond yield), respectively. Mankiw, Reis, and Wolfers (2004) and Ball and Croushore (2003) employ models similar to (4). They regress forecast errors on the variables introduced in our set of regressors. However, our model features past errors and changes in the relevant regressors as determinants of the current forecast error. Evidence of serial correlation in the forecast error process indicates that there is inefficient exploitation of information from last year's forecast. In this case the RE hypothesis is violated. Figure 2 reports the total R^2 for each regression as well as the contribution of each regressor to the explanation of the variation of a dependent variable (Scherrer, 1984).[18] Table A2 and Figure A1 in the appendix report the estimated coefficients.

Insert Figure 2 about here

It turns out that the coefficients associated with the horizontal spread and the cycle indicator are never significantly different from zero. Below this level, the cycle indicator exerts a negative effect. The response associated with the last observed forecast error is fairly constant up to the 30^{th} percentile (see Figure A1c), declining thereafter and then assuming a U-shaped pattern, with a minimum occurring around the at the 55^{th} percentile. On a priori grounds, only the error of the professional forecasters is expected to be positive and significantly different from zero to confirm rationality. Our results show that the response is first constant, then hump-shaped around the 55^{th} percentile, while it decreases in the last deciles. As to the effect of $\Delta\pi_t$, our estimates show that the response increases monotonically from the 51^{st} percentile onwards, thus displaying a substantial degree of overreaction to changes in current inflation.

[17] This equation could also be considered as a test of rationality. The test would be based on the null hypothesis that $\alpha = \beta = \gamma = 0$. To assess Carroll's (2003a, 2003b) finding that the transmission effect from professional forecasters to households is quite slow, we also add several lags of the SPF. However, these turn out to have no explanatory power.

[18] The coefficient of multiple determination measures the proportion of the variance of a dependent variable y explained by a set of explanatory variables. It can be computed as $R^2 = \sum_{j=1}^{k} a_j r_{yx_j}$, where a_j is the standardized regression coefficient of the j^{th} explanatory variable and r_{yx_j} is the simple correlation coefficient (Pearson's r) between y and x_j. Scherrer defines $a_j r_{yx_j}$ as the contribution of the j^{th} variable to the explanation of the variance of y.

The coefficient of determination declines as we move towards the upper end of the distribution. Nevertheless, it does not follow a monotonic pattern, but displays a marked hump-shaped behavior in the middle forecast range and a U-shaped pattern from the 70^{th} percentile onward. It appears that the last observed error captures substantial variance in the LHS forecasts, which display a market degree of backward lookingness. Forecasts in this range do not rely on current inflation. The variance of the forecast error on the RHS is almost exclusively explained by the variance of the change in actual inflation. This is a further signal of the pessimism reflected in the overreaction of these agents to changes in contemporaneous inflation.

In the central forecast range the contribution of past errors decreases, while the contribution of the SPF error acquires further importance. Regarding professional forecasters as rational agents, we can actually infer that the middle range is the least biased, especially around the $50^{th} - 55^{th}$ percentile. In this region the error of professional forecasters is almost the only relevant variable to explain the forecast error.

3.3. Adaptive Expectations. In this section we analyze the degree of adaptiveness in households' inflation forecasts. The idea of adaptive expectations originated in Fisher (1930) and was formally introduced in the 1950s by several authors, e.g. Fisher (1930). Nerlove, Grether, and Carvalho (1979) were the first to model expectations as an autoregressive process and labelled them as quasi-rational expectations. The following regression model, conceived as a preliminary assessment of the degree of adaptiveness, is equivalent to an adaptive expectations formula:

$$\pi^k_{t|t-12} = \pi^k_{t-13|t-25} + \zeta \left(\pi_{t-13} - \pi^k_{t-13|t-25} \right) + \varepsilon_t, \tag{5}$$

Under this rule, agents revise their expectations according to the last observed forecast error. Parameter ζ is labelled as the "error-adjustment" coefficient. It captures the speed of adjustment of present forecasts to past forecast errors. As interviewees are asked to forecast inflation over the next year (hence they make their forecast at time $t-12$), the revision will be based on the previous period's forecast, which has been carried out at time $t-25$. A word of caution is in order at this stage. In the adaptive learning approach, which is discussed in further detail in the next section, adaptive behavior reflects in the estimation of the parameters of the Perceived Law of Motion (PLM). The adjustment of these parameters towards the value consistent with the REE depends on past forecast errors, while adaptive expectations postulate that agents revise their current expectation based on past forecast errors. In this case the error-adjustment coefficient is assumed to be constant.

Results. In Figures 3(a)-(b) we plot, for each percentile, the gain parameter and the corresponding R^2 obtained from (5).

Insert Figure 3 about here

Overall, forecasts on the RHS of the median forecast at least partly behave in an adaptive manner, while past errors have little or no explanatory power for LHS forecasts.[19] As to the

[19] A negative estimated gain on the LHS could reflect a divergent behavior, as errors become larger over time. As displayed in Figure 3(b), this model does not provide a good fit for the LHS of the median forecast.

estimated constant gain and the R^2, we observe a clear hump-shaped response between the 40^{th} and 99^{th} percentile, with a peak occurring at the 75^{th} percentile.

3.4. Sticky Information. Carroll (2003a, 2003b) designs an epidemiological framework to study how the Michigan Survey respondents form their expectations. He models the evolution of inflationary expectations based on the assumption that households update their information set from news reports, which in turn are influenced by the expectations of professional forecasters. His results suggest that the diffusion process is slow, due to households' inattentiveness. Moreover, SPF inflation expectations are found to Granger-cause households' inflation expectations, whereas the opposite does not hold true.

Testing for Sticky Information – Static Case. We estimate a simple regression in the vein of Carroll (2003a):

$$\pi^k_{t|t-12} = \lambda_1 \pi^F_{t|t-12} + (1 - \lambda_1)\,\pi^k_{t-1|t-13} + \varepsilon_t. \tag{6}$$

As Carroll (2003a) points out, news about inflation spread slowly across agents, reaching only a fraction λ_1 of the population in each period. The model is estimated under the assumption that coefficients sum up to 1, although this restriction is not likely to be satisfied across all percentiles.[20]

We also find evidence of time-varying degrees of heterogeneity in the frequency of information updating over the cross-sectional range of responses. We consider two subsamples of forecasts, namely pre-1988 and post-1988, so that two different inflationary regimes are considered. On the one hand, our results suggest that forecasts slightly below the median entail a higher degree of unbiasedness in the post-1988 subsample, as they reflect higher frequency of information updating.[21] On the other hand, forecasts slightly above the median reflect systematic errors when inflation is lower and more stable, thus supporting the inattentiveness argument. This is explored further in the next subsection.

Results. Figure 4(a) plots, for each percentile, the estimated λ_1^{-1}. This provides us with a measure of the average updating period. The estimation confirms the existence of static behavior in the informational structure up to the 40^{th} percentile. From this point up to the 91^{st} percentile, a U-shaped pattern emerges, with a minimum occurring at the 50^{th} percentile. This translates into an average minimum updating period of 7 months. Carroll (2003a) finds similar results: his estimate of λ_1^{-1} is 11 months for the mean[22], while Döpke, Dovern, Fritsche, and Slacalek (2006a) estimate an average updating period of 18 months for the Euro area. Mankiw, Reis, and Wolfers (2004) and Branch (2007) set λ at 0.1, which for monthly data implies an

[20] It should be pointed out that this model is derived under the assumption that: (i) inflation follows a random walk process; (ii) $\pi^k_{t|t-13} \approx \pi^k_{t-1|t-13}$ (see Döpke et al., 2006a).

[21] On the one hand, one could interpret these results as a critique to our approach which does not allow agents to switch across different percentiles. On the other hand, our results point out that the intervals identified above preserve their main characteristics, both in terms of information and in terms of expectation formation.

[22] Mankiw and Reis (2002) have implemented $\lambda = 0.25$ (average updating 12 months) in their simulations, assuming quarterly data.

average frequency of 10 months. Branch (2007) further investigates the sticky information argument by allowing for switching between different updating frequencies.

Testing for Sticky Information - State Dependent Coefficients. When inflation matters agents update their information set more frequently, in order to produce more accurate forecasts. In addition, in periods of marked macroeconomic turmoil the amount of media coverage is generally higher, hence the cost of acquiring information is lower. We assume that a higher proportion of agents pays attention to new information when inflation is higher, as the opportunity cost of being inattentive is significantly higher during these phases. To test this hypothesis, we relax the assumption of linearity in equation (6). We assume a non-linear structure in the form of a logistic smooth-transition autoregressive (LSTAR) model:[23]

$$\pi^k_{t|t-12} = \lambda_1 F\left(\pi_{t-12}\right)\pi^F_{t|t-12} + \left[1 - \lambda_1 F\left(\pi_{t-12}\right)\right]\pi^k_{t-1|t-13} + \varepsilon_t, \tag{7}$$

where F indicates the following logistic function:[24]

$$F\left(\pi_{t-12}\right) = \frac{1}{1 + \exp\left[-\upsilon\left(\pi_{t-12} - c\right)\right]}, \tag{8}$$

where υ can be interpreted as a parameter measuring the speed of responsiveness, whereas c is a threshold coefficient. The approach consists of estimating λ_1 by means of least squares while running a grid search on υ and c, in order to find the combination of values that minimizes the SSE for each percentile.

Insert Figure 4 about here

Results. We estimate positive coefficients in the transition function at every percentiles. Since the SSE is always lower under (7) we can assert that the non-linear version outperforms the linear version.[25] Responses between the 59^{th} and 79^{th} percentile are clearly associated with the inattentiveness argument. Within this range a higher level of attentiveness is displayed in periods of high inflation compared to periods of low inflation.

Figure 4(b) reports the estimated $\left[\lambda_1 F\left(\pi_{t-12}\right)\right]^{-1}$, which is a time-varying estimate of the average updating period for the 52^{nd} and 63^{rd} percentile. As we can notice, the average updating period for the 63^{rd} percentile behaves in accordance with the inattentiveness view. At the beginning of the sample, the average updating period is rather low, as inflation is higher, as the opportunity cost of not updating the information set. The optimal coefficients in the transition function for this percentile are $\upsilon = 0.21$ and $c = 2.58$. The latter can be interpreted as a perceived implicit inflation target of the FED. The dynamics of the average updating period for the 52^{nd} percentile is quite different. This percentile displays lower attentiveness only sporadically. The optimal coefficients in the transition function for the 52^{nd} percentile are $\upsilon = 3.18$ and $c = 7.40$.

[23] For details about smooth-transition regression models see Granger and Teräsvirta (1993).

[24] We have also tried different forms of transition function, but the symmetric case we present outperforms the other alternatives in terms of SSE.

[25] The average difference in the SSE is about 0.595.

The interpretation of these coefficients is different from the previous case, as υ is higher than 1. Consequently, c cannot be interpreted as a perceived inflation target. For these agents, the difference between the linear and non-linear model is negligible. The $50^{th} - 58^{th}$ percentile range exhibits a similar response, although for higher percentiles in this range the variability of the estimated average updating frequency is higher. Thus, when inflation is low higher inattentiveness is observed. Analogous evidence applies above the 80^{th} percentile, although this area is associated with a much higher average time to update information.

Model (7) can also be interpreted as an alternative specification to the one proposed by Branch (2007). In this case the choice between different updating frequencies is modelled through a mechanism á la Brock and Hommes (1997). Results indicate that the majority of agents update their information set every $3 - 6$ months, while fewer agents update their information set every month. Some agents update every 9 months or even less frequently. Our results stand in partial contrast to Branch (2007). We provide evidence that information updating is less frequent on the LHS of the median forecast. RHS forecasts are formed in accordance with the inattentiveness argument and generally display lower updating frequency in periods of stable inflation (approximately every two years). Conversely, the frequency of information updating increases when inflation is higher.

3.5. Adaptive Learning. This section is designed to assess the empirical significance of adaptive learning in the MSHE distribution. Different learning rules are considered to test for convergence to rational expectations (perfect foresight) and to measure the speed of learning. For a comprehensive discussion on different learning rules and convergence to rational expectations see Evans and Honkapohja (2001).

The adaptive expectations model (5) has been designed to provide a preliminary assessment of the degree of adaptiveness in the data. In the adaptive learning literature, it is assumed that agents behave like econometricians, using the available information at the time of the forecast. Let us assume that the forecasters considers the following perceived law of motion (PLM):

$$\pi_{t|t-12} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-13} + \varepsilon_t, \tag{9}$$

whose coefficients are assumed to be time-varying and follow a specific updating mechanism that will be briefly detailed. When agents estimate their PLM, they exploit the available information up to period $t-1$. As new data become available, they update their estimates according to a constant gain learning (CGL) rule or a decreasing gain learning (DGL) rule. First, we focus on stochastic gradient learning and then on least squares learning, both under constant gain (CG) or decreasing gain (DG). Let X_t and $\widehat{\phi}_t$ be the following vectors: $X_t = \begin{pmatrix} 1 & \pi_t \end{pmatrix}$ and $\widehat{\phi}_t = \begin{pmatrix} \phi_{0,t} & \phi_{1,t} \end{pmatrix}'$. When relying on stochastic gradient learning, agents update coefficients according to the following rule (see Evans, Honkapohja and Williams, 2005):

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-25}\left(\pi_{t-12} - X_{t-25}\widehat{\phi}_{t-13}\right). \tag{10}$$

In the updating algorithm for DGL, we replace ϑ with $\frac{\iota}{t}$. When using least squares learning, agents also take into account the matrix of second moments of X_t, R_t. Under CGL, coefficients are updated according to:

$$\begin{aligned}\widehat{\phi}_t &= \widehat{\phi}_{t-1} + \vartheta R_{t-1}^{-1} X'_{t-25}\left(\pi_{t-12} - X_{t-25}\widehat{\phi}_{t-13}\right); \qquad (11)\\ R_t &= R_{t-1} + \vartheta\left(X_{t-25}X'_{t-25} - R_{t-1}\right). \qquad (12)\end{aligned}$$

Alternatively, when specifying the updating algorithm under decreasing gain learning we simply replace ϑ with $\frac{\iota}{t}$.

A Standard Updating Mechanism. In order to implement the adaptive learning approach to MSHE data, we specify the following PLM:

$$\pi^s_{t|t-12} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-13} + \varepsilon_t, \qquad (13)$$

where superscript "s" stands for simulated forecast. Our exercise is designed to search for the combination of initial values and gain parameter that replicates each percentile as closely as possible. The drawback implicit in this approach is that initial values of $\widehat{\phi}_t$ for 12 periods have to be assumed. In the recursive estimation of the gain parameter setting initial values represent the main problem: this is extensively discussed in Carceles-Poveda and Giannitsarou (2007). Previous estimations of models under adaptive learning (e.g. Milani, 2005) have generally split the time series into two subsamples. Thus, the first subsample is only used to estimate the set of initial values for the parameters in the PLM which are then employed for the recursive estimation of the gain parameter in the second subsample. Clearly, the main practical drawback of this approach is that it does not allow the researcher to fully exploit the data available. Moreover, this approach does not abstract from the risk that learning dynamics could just result as a statistical artifact due to a non-optimal initialization.

Our approach abstracts from this criticism, as we search for the optimal combination of the initial values and the gain parameter, thus preserving the sample structure and optimizing the initialization procedure. In practice, several forecast series ($\pi^s_{t|t-12}$) are simulated, by means of a multidimensional grid search, under different combinations of ϑ and $\widehat{\phi}$. We then select the gain parameter ϑ (or ι under DGL) and the set of initial values $\widehat{\phi}$ that minimize the sum of squared errors (SSE), i.e. $\left(\pi^s_{t|t-12} - \pi^k_{t|t-12}\right)^2$.[26] This strategy can also be regarded as a test for learning dynamics. If the gain is found to be positive under this method of initialization, then the series would exhibit learning for all other initialization methods with a higher (or equal) gain.

Results. The 65^{th}-98^{th} percentile range displays evidence in line with CG gradient learning. The estimated gain, reported in Figure 5(a), displays a hump over this forecast range, with a peak at 2.1×10^{-4}. This maximum is located between the 71^{st} and 73^{rd} percentile. The

[26] However, this approach has an obvious practical inconvenience, as running a grid search on several variables is computationally very intensive.

DG version of gradient learning turns out to be significant for the $70^{th} - 96^{th}$ percentile range. Within this range, the estimated gain displays properties similar to those detected under CGL. Both CGL and DGL exhibit a second minor hump in the RHS of the distribution. This is more pronounced under DGL [see Figure 5(a)]. The highest gain is estimated around the $76^{th} - 77^{th}$ percentile $(0.007t^{-1})$. To compare both versions of gradient learning, we plot their SSEs in Figure 5(b). Our results suggest that the CG version of gradient learning generally provides a better description of agents' behavior, especially around the 70^{th} percentile.

Orphanides and Williams (2005a) suggest a value of the gain coefficient between 0.01 and 0.04, whereas Milani (2007) estimates a gain of 0.0183. These estimates are obtained from quarterly data. An estimated gain of 0.02 means that agents rely on 12.5 years of data to produce their forecast. As in this study we explore monthly data, an estimate of 2.1×10^{-4} implies that roughly 400 years of data are employed to produce a forecast. However, our estimates should only be regarded as the lower bound of the gain coefficient for the reasons exposed above. Furthermore, Eusepi and Preston (2008) suggest that the gain might be comprised between 0.0015 and 0.0029. This value is closer to our estimates.

Insert Figure 5 about here

When taking into account the matrix of second moments,[27] we find very similar results, as covariance terms are found to be rather small. We also consider alternative PLMs. In the following formulation previous period inflation is replaced by previous period forecast:

$$\pi^s_{t|t-12} = \phi_{0,t-1} + \phi_{1,t-1}\pi^k_{t-1|t-13} + \varepsilon_t. \tag{14}$$

This formulation is found to provide a better fit compared to (13) [see Figure 5(d), where we compare the SSE under the two updating algorithms]. Moreover, in this case some learning dynamics can also be detected on the LHS of the distribution. The $1^{st} - 9^{th}$ and the $63^{rd} - 99^{th}$ percentile range display adaptive behavior consistent with CGL dynamics. We obtain similar results under DGL, as evidence of learning is detected for the $1^{st} - 9^{th}$ and the $69^{th} - 99^{th}$ percentile range. In the CG case, the response pattern on the RHS is bell-shaped, with the highest gain occurring at the $78^{th} - 79^{th}$ percentile $\left(5.5 \times 10^{-5}\right)$. The response under DG is also hump-shaped on the RHS, reaching the highest gain at the $75^{th} - 76^{th}$ percentile $\left(0.0067t^{-1}\right)$.

An Iterative Representation of the PLM. We now introduce a PLM featuring the last observed rate of inflation:

$$\pi^s_{t|t-1} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1} + \varepsilon_t. \tag{15}$$

We implement the following gradient learning updating algorithm:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1}\left(\pi_t - X_{t-1}\widehat{\phi}_{t-1}\right). \tag{16}$$

[27] We set this matrix to be constant and equal to the sample average. In this case RLS estimates are approximately linear combinations of the gains obtained under stochastic gradient learning, as covariance terms are quite small.

As we consider 12 months ahead forecasts, agents are assumed to implement the following rule:

$$\pi^s_{t+12|t} = \phi_{0,t-1}\left[1 + \phi_{1,t-1} + \left(\phi_{1,t-1}\right)^2 + ... + \left(\phi_{1,t-1}\right)^{12}\right] + \left(\phi_{1,t-1}\right)^{13}\pi_{t-1}. \tag{17}$$

The advantage of this approach is that we only need to assume initial values for 1 period.

Results. This recursive algorithm delivers results analogous to those obtained under (9). However, this learning method provides a slightly less accurate explanation of inflation forecasts. Our estimates suggest that the 65^{th}-99^{th} percentile range displays CGL dynamics. In this case, optimal gains exhibit a marked hump-shaped pattern. The peak occurs at $2.35 * 10^{-4}$, between the 79^{th} and 82^{nd} percentile [see Figure 5(e)].In the DG case, the gain peaks at $0.0125t^{-1}$, between the 74^{th} and 75^{th} percentile. As in the previous version, CGL constantly outperforms DGL [see Figure 5(f)].

An Updating Mechanism Based on Expected Future Errors. In the next updating mechanism we allow for a higher degree of forward lookingness compared to the one traditionally assumed in the adaptive learning literature. We introduce a novel mechanism of expectation formation which presumes that agents update their coefficient estimates with respect to new information about future inflation. In this case, new information is proxied by SPF forecasts. Implicitly, this model states that agents update their information set from the media, which are assumed to transmit the expectations of the professional forecasters. The underlying mechanism is consistent with the epidemiological view advanced by Carroll (2003a), and implicitly represents a combination of adaptive learning and sticky information. We assume a PLM of the following form:

$$\pi^s_{t+12|t} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1} + \varepsilon_t. \tag{18}$$

The following gradient learning updating algorithm is considered:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1}\left(\pi^F_{t+12|t} - X_{t-1}\widehat{\phi}_{t-1}\right). \tag{19}$$

Results. Our results suggest that learning dynamics is displayed above the 52^{nd} percentile under CG and the 51^{st} percentile under DG. The highest gain is $7.40 * 10^{-4}$ under CG and $0.0200t^{-1}$ under DG. As we can observe in Figure 6(b), CGL significantly outperforms DGL after the 65^{th} percentile.

Insert Figure 6 about here

Interestingly, this approach produces a cross-sectional pattern of the gain parameter which is more in line with what we should expect on a priori grounds. The highest gain is reached slightly above the median and declines thereafter. Also, compared to the previous updating algorithms, a wider proportion of forecasts are consistent with (19). This signals that, despite the fact updating dynamics can be detected, substantial forward lookingness characterizes the updating procedure. Moreover, as it will emerge in the next updating rule, the distribution of forecasts reflects a higher informational content compared the one proxied by the prediction of the professional forecasters.

An Updating Mechanism Based on Future Errors. Consistent with the view advanced above, we allow for the possibility that agents access more information about future developments of inflation compared to what reflected in SPF forecasts. One motivation for this concern is that several studies have documented the presence of herding behavior in the predictions of professional forecasters.[28] We proxy this information with next period's inflation. The PLM reads as in (18). We implement a gradient learning updating algorithm:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1}\left(\pi_{t+12} - X_{t-1}\widehat{\phi}_{t-1}\right). \tag{20}$$

We also consider a least squares learning version of (20).

Insert Table 1 and Figure 7 about here

Results. The forward-looking updating mechanism (20) allows us to assess the importance of learning from new information, compared to previous versions characterized by a backward-looking perspective. Our results suggest that RHS forecasts can be associated with this version of adaptive learning. In order to support this evidence, we also explore a wider set of potential PLMs. We start with (13). Results suggest that data display learning dynamics from the 55^{th} percentile under CG and the 56^{th} percentile under DG. In both cases, the gain immediately jumps to the highest value and decreasing thereafter. The highest gain is estimated at $1.125 * 10^{-3}$ and the lowest SSE is reached at the 68^{th} percentile [see Figures 7(a)-(b)]. Compared to estimates obtained under the first version, this gain can be regarded as more realistic, as it suggests that about 74 years of data are used to produce forecasts. Nonetheless, this estimate is still quite high. Under DGL, the highest gain is $0.0445t^{-1}$ while the SSEs are very similar to those obtained under CGL. Strictly speaking, CGL performs slightly better for most of the percentiles, except for those between the 63^{rd} and 69^{th} percentile.

As to least squares learning, we set the variance-covariance matrix in line with the sample average. The results in this case are very similar to those obtained under stochastic gradient learning. The maximum optimal gain is $8.5 * 10^{-8}$ under CGL and $3.5 * 10^{-6}t^{-1}$ under DGL.

We also explore learning with PLMs that alternatively feature the second lag of inflation, output gap and SPF inflation forecasts. We find that the PLM implementing SPF inflation forecasts performs better compared to other options, especially under DGL [see Figure 7(d)]. The pattern of the optimal gain is quite similar across competing PLMs. These PLMs indicate that agents between the 54^{th} and 98^{th} percentile behave in accordance to adaptive learning based on SPF forecasts. Figure 7(f) plots the SSEs, whereas Table 1 reports the maximum gains. The optimal gain under CGL is estimated between 0 and 0.051. In addition, we can claim that DGL provides a better fit.

[28] This effect, studied by Scharfstein and Stein (1990), Banerjee (1992) and Zwiebel (1995), is based on the people who make forecasts occasionally being afraid of deviating from the majority or consensus opinion. Pons-Novell (2003) documents this fact empirically.

Overall, these results confirm that forecasts on the RHS are more in line with this version of learning dynamics, compared to classical models of learning.

4. Discussion

Our analysis highlights the presence of a marked degree of heterogeneity in the process of expectation formation. Our results allow us to identify three regions of the distribution that correspond to different underlying mechanisms of expectation formation. On the one hand, we can consider the interval on the LHS of the distribution as the one characterized by forecasts that do not exploit the relevant information. On the other hand, predictions on the RHS of the median forecast reflect significant overreaction to information about future inflation. Intuitively, forecasts in the middle range of the density are unbiased. Table 2 reports, for each range of responses, the models of expectation formation that are consistent with the data. Moreover, we report the variables exploited in the prediction of future inflation for each of these models and the degree of reliance on these variables (partial reliance $=P$; full reliance$=F$; over-reaction$=O$).

Insert Table 2 about here

LHS forecasts display a substantial degree of backward lookingness. As shown in Table 2, this forecast range can be divided into three further sub-intervals. In the first sub-interval (up to the 10^{th} percentile), forecasts are nearly static, as the information set is virtually never updated. Past inflation is not taken into account. Only past forecasts are considered and, to some extent, the cycle indicator. Moreover, we find some support for adaptive learning, where parameters are updated with respect to past errors. To conclude, forecasts in this sub-interval mainly display some form of AR(1) rule and, from time to time, coefficients are updated with respect to the last observed error. A second sub-interval on the LHS can be identified between the 11^{th} and 30^{th} percentile. Also in this range forecasts do not reflect any systematic information updating. Compared to the previous sub-interval, past inflation in the PLM is now significant. No form of adaptive behavior is significant. We could characterize this sub-interval through a PLM featuring an intercept, past forecasts and past inflation. The third sub-interval incorporates forecasts lying between the 31^{th} and 49^{th} percentile. Information updating occurs on a more regular basis, especially after the 40^{th} percentile. Past inflation is fully exploited into the PLM, as well as information on future developments in inflation dynamics through the predictions of professional forecasters. In this interval, the dependence of the forecast error on past errors gradually decreases as we move toward the RHS. The dynamics of the percentiles in this forecast range could be characterized by a PLM featuring the intercept, past forecast, past inflation and SPF forecasts.

Forecasts in the central range of the empirical distribution are generally unbiased. Moreover, the $\left[50^{th}, 55^{th}\right]$ percentile interval displays regular information updating. As expected, the SPF forecast error is the only explanatory variable for errors in this range.

The RHS of the distribution displays forecasts in line with theories of adaptive learning and inattentiveness. Predictions above the 56^{th} percentile can be further divided into four sub-intervals. The first sub-interval can be identified between the 56^{th} and 66^{th} percentile. Forecasts comprised in this range of responses display inattentiveness. The average updating frequency ranges between 8 months and 30 months. Adaptive learning is also a plausible explanation of the process of expectation formation in this forecast range, as coefficients in the PLM are updated with respect to new information. Some degree of pessimism characterizes these forecasts, as prediction errors are increasingly associated with changes in the current inflation as we move towards the RHS. The second sub-interval on the RHS of the distribution encompasses forecasts between the 67^{th} and 72^{nd} percentile. This region of the empirical density behaves in line with a version of adaptive learning featuring updating with respect to new information (SPF in the PLM). Forecasts between the 73^{rd} and 90^{th} percentile can be grouped in the third sub-interval on the RHS. Their dynamics can be replicated under a CGL algorithm with a PLM featuring past forecasts. Coefficients in the PLM are updated with respect to the last observed error. Forecasts in this group reflect information updating at a lower frequency, compared to what detected in the previous regions of the empirical density. Moreover, changes in actual inflation acquire a predominant role in the determination of forecast errors in this area. The last range of forecasts on the RHS can be placed above the 91^{st} percentile. This is associated with a DG version of learning, where coefficients are updated with respect to new information. Information is updated rather infrequently and forecast errors are almost exclusively explained by changes in current inflation.

Concluding Remarks

This paper deals with the development of techniques for the empirical analysis of adaptive learning and information stickiness. These methodologies are applied to the distribution of households' inflation expectations collected by the University of Michigan Survey Research Center. In order to account for the degree of asymmetry and multimodality in the empirical density, we apply our techniques to the entire cross-sectional range of forecasts.

First, we extend the epidemiological framework proposed by Carroll (2003a) to account for the possibility that agents are more likely to update their information set on a regular basis in periods of high inflation. The resulting LSTAR model provides a reasonable description of the forecasts range in the upper end of the distribution of inflation forecasts. This region displays greater attentiveness in periods of high and volatile inflation.

We then introduce a novel technique to detect adaptive learning in the distribution of forecasts. We tackle the problem of initializing the learning algorithm and propose a computational technique to search for the optimal combination of initial values (for the parameters of the PLM) and gain parameter, thus preserving the sample structure and optimizing the initialization procedure. This procedure allows us to fully exploit the data available and to avoid that the detection of learning dynamics results as a mere statistical artifact due to a non-optimal initialization. We also propose an alternative mechanism of expectation formation, whereby households are assumed to update their forecast with respect to (expected) future errors, which are reflected in the difference between their forecasts and the predictions of the professional forecasters. This model draws on the epidemiological view advanced by Carroll (2003a), and represents a combination of adaptive learning and sticky information.

The implementation of these techniques generates a set of stylized facts that allows us to identify three regions of the distribution that correspond to different underlying mechanisms of expectation formation: a static or highly autoregressive region on the left hand side of the median, a nearly rational region around the median, and a fraction of forecasts on the right hand side of the median forecast produced in accordance with adaptive learning and sticky information.

References

Bakhshi, H., and A. Yates (1998): "Are U.K. Inflation Expectations Rational?," Discussion Paper 81, Bank of England Working Paper.

Ball, L., and D. Croushore (2003): "Expectations and the Effects of Monetary Policy," *Journal of Money, Credit and Banking*, 35(4), 473–484.

Banerjee, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107(3), 797–817.

Bernanke, B. (2007): "Inflation Expectations and Inflation Forecasting," Discussion paper, Speech at National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts.

Branch, W. A. (2004): "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations," *Economic Journal*, 114(497), 592–621.

——— (2007): "Sticky information and model uncertainty in survey data on inflation expectations," *Journal of Economic Dynamics and Control*, 31(1), 245–276.

Branch, W. A., and B. McGough (2006): "A New Keynesian Model with Heterogeneous Expectations," Mimeo, University of California, Irvine.

Brock, W. A., and C. H. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65(5), 1059–1096.

Carceles-Poveda, E., and C. Giannitsarou (2007): "Adaptive learning in practice," *Journal of Economic Dynamics and Control*, 31, in press.

Carroll, C. D. (2003a): "Macroeconomic Expectations Of Households and Professional Forecasters," *The Quarterly Journal of Economics*, 118(1), 269–298.

——— (2003b): "The Epidemiology of Macroeconomic Expectations," NBER Working Papers 8695, National Bureau of Economic Research, Inc.

Croushore, D. (1998): "Evaluating inflation forecasts," Working Papers 98-14, Federal Reserve Bank of Philadelphia.

Curtin, R. (1996): "Procedure to estimate price expectations," Mimeo, University of Michigan.

——— (2005): "Inflation Expectations: Theoretical Models and Empirical Tests," Mimeo, University of Michigan.

Döpke, J., J. Dovern, U. Fritsche, and J. Slacalek (2006a): "The Dynamics of European Inflation Expectations," Discussion Papers of DIW Berlin 571, DIW Berlin, German Institute for Economic Research.

Eusepi, S., and B. Preston (2008): "Expectations, Learning and Business Cycle Fluctuations," Working Paper 14181, National Bureau of Economic Research.

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

Evans, G. W., S. Honkapohja, and N. Williams (2005): "Generalized Stochastic Gradient Learning," NBER Technical Working Papers 0317, National Bureau of Economic Research.

Fisher, I. (1930): *Theory of Interest.* Macmillan, New York.

Granger, C. W. J., and T. Teräsvirta (1993): *Modelling Nonlinear Economic Relationships.* Oxford University Press, New York.

Jonung, L., and D. Laidler (1988): "Are Perceptions of Inflation Rational? Some Evidence for Sweden," *American Economic Review*, 78(5), 1080–87.

Kuttner, K. N. (1994): "Estimating Potential Output as a Latent Variable," *Journal of Business of Economic Statistics*, 12(3), 361–68.

Mankiw, N. G., and R. Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 117(4), 1295–1328.

Mankiw, N. G., R. Reis, and J. Wolfers (2004): "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual 2003*, 18, 209–248.

Milani, F. (2005): "Adaptive Learning and Inflation Persistence," Macroeconomics 0506013, EconWPA.

——— (2007): "Expectations, learning and macroeconomic persistence," *Journal of Monetary Economics*, 54(7), 2065–2082.

Nerlove, M. L., D. M. Grether, and J. L. Carvalho (1979): *Analysis of Economic Time Series: A Synthesis.* New York: Academic Press, Inc, revised edition, 1995 edn.

Orphanides, A., and J. C. Williams (2003): "Imperfect Knowledge, Inflation Expectations, and Monetary Policy," NBER Working Papers 9884, National Bureau of Economic Research, Inc.

——— (2005a): "The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations," *Journal of Economic Dynamics and Control*, 29(11), 1927–1950.

——— (2005b): "Inflation scares and forecast-based monetary policy," *Review of Economic Dynamics*, 8(2), 498–527.

Pesaran, M. H. (1987): *The Limits to Rational Expectations.* Basil Blackwell, Oxford, reprinted with corrections 1989 edn.

Pfajfar, D., and E. Santoro (2008a): "Asymmetries in Inflation Expectation Formation Across Demographic Groups," Cambridge Working Papers in Economics 0824, Faculty of Economics, University of Cambridge.

Pons-Novell, J. (2003): "Strategic bias, herding behaviour and economic forecasts," *Journal of Forecasting*, 22(1), 67–77.

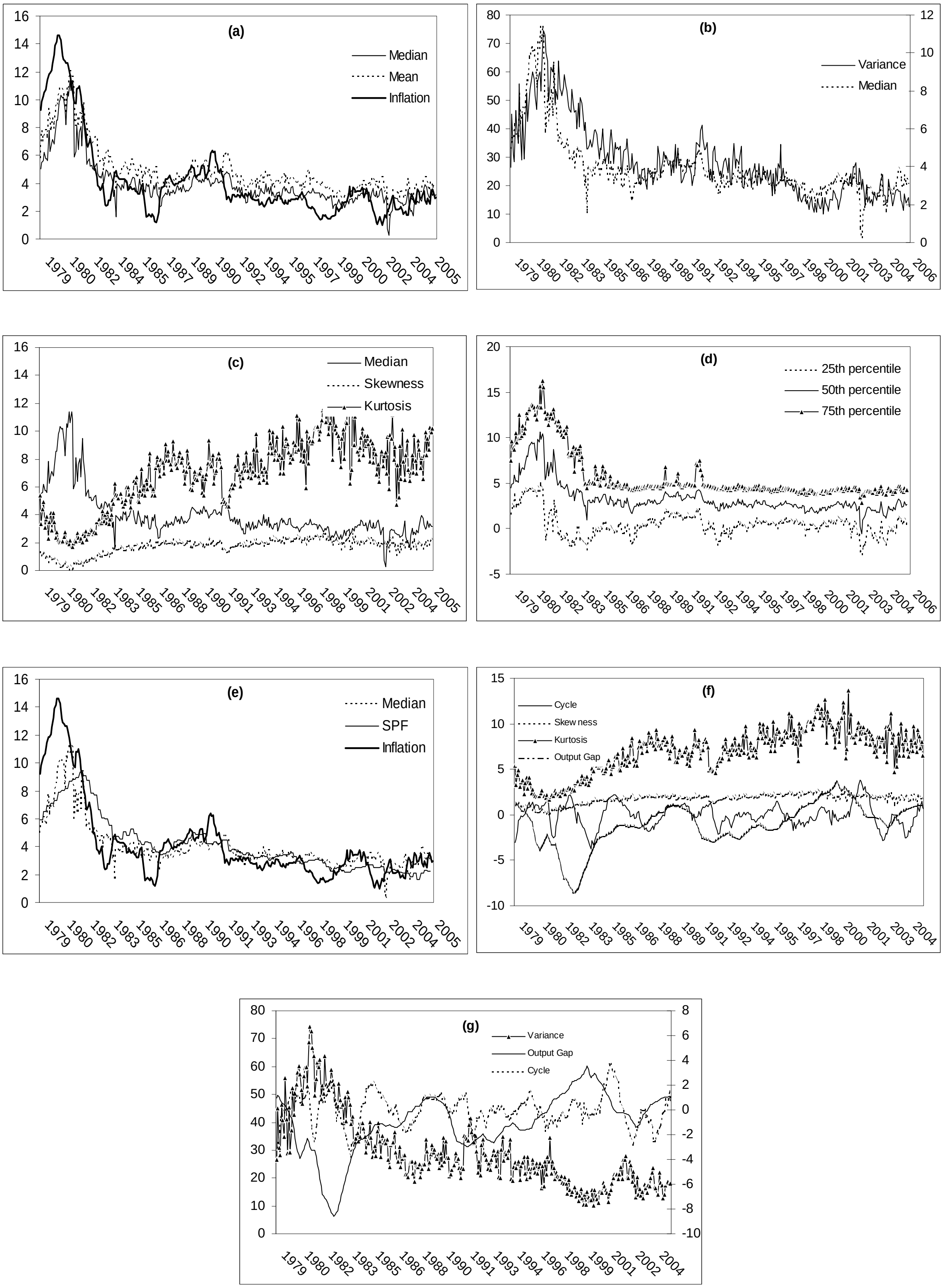
Roberts, J. M. (1997): "Is inflation sticky?," *Journal of Monetary Economics*, 39(2), 173–196.

Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465–79.

Scherrer, B. (1984): *Biostatistique*. Gaëtan Morin, Chicoutimi.

Zwiebel, J. (1995): "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 103(1), 1–25.

5. Figures and Tables



Figures 1(a)-1(g): Empirical moments of the MSHE distribution (realized date).

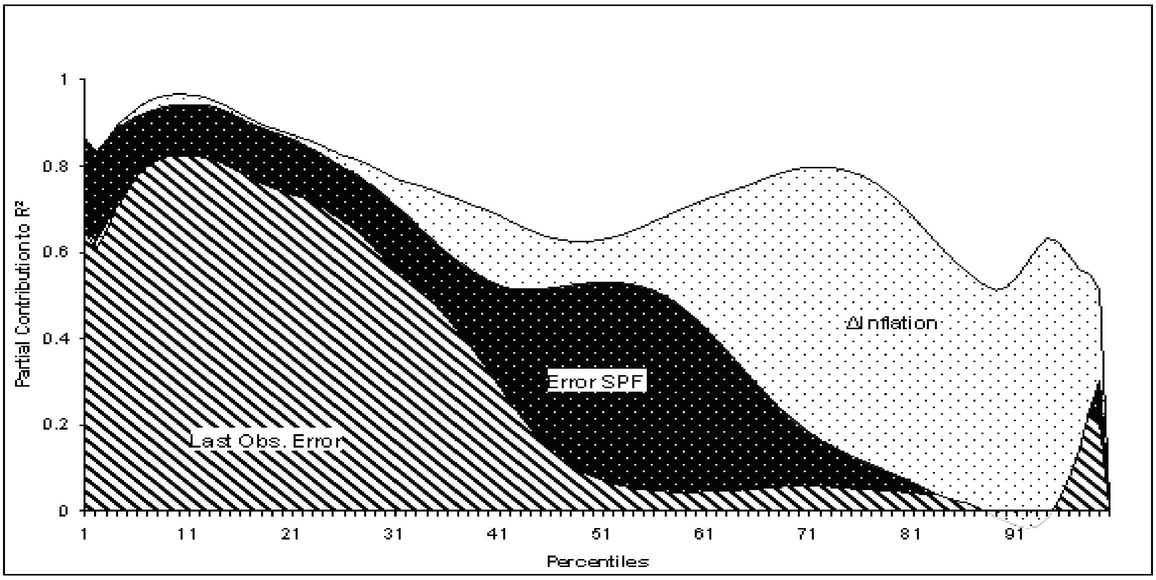
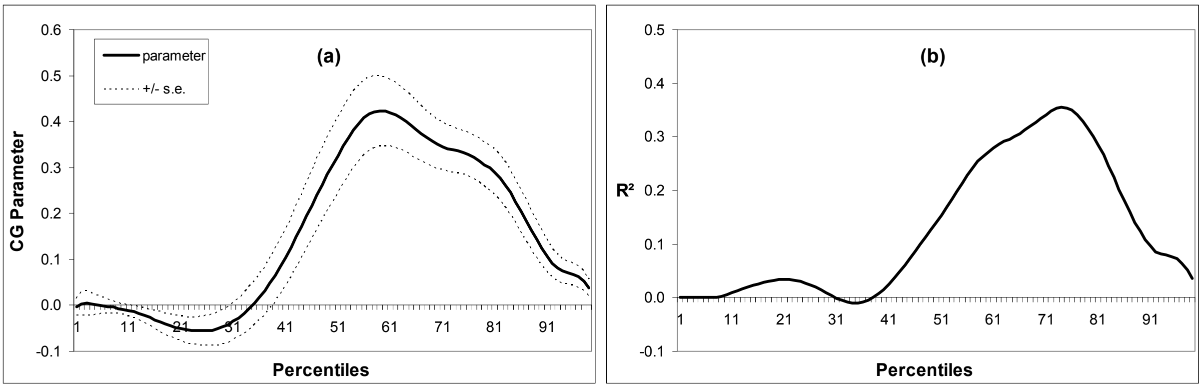
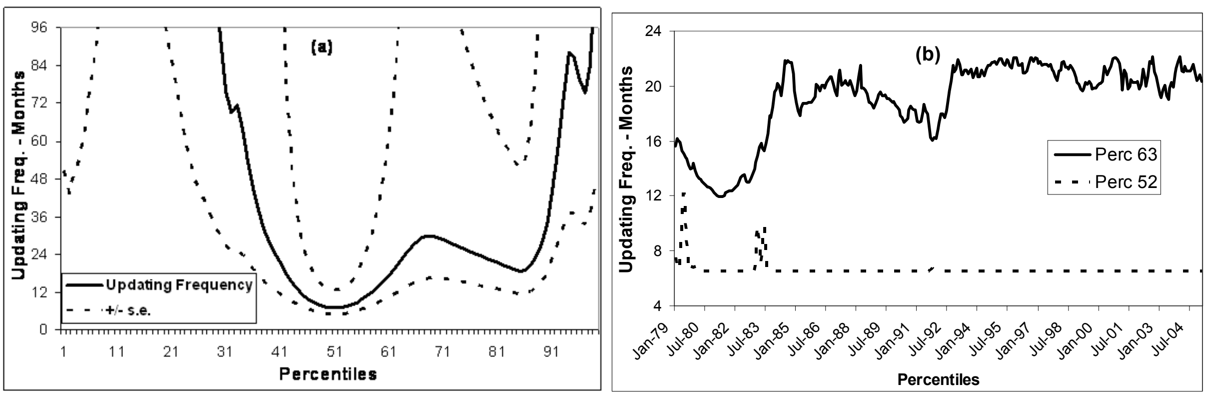


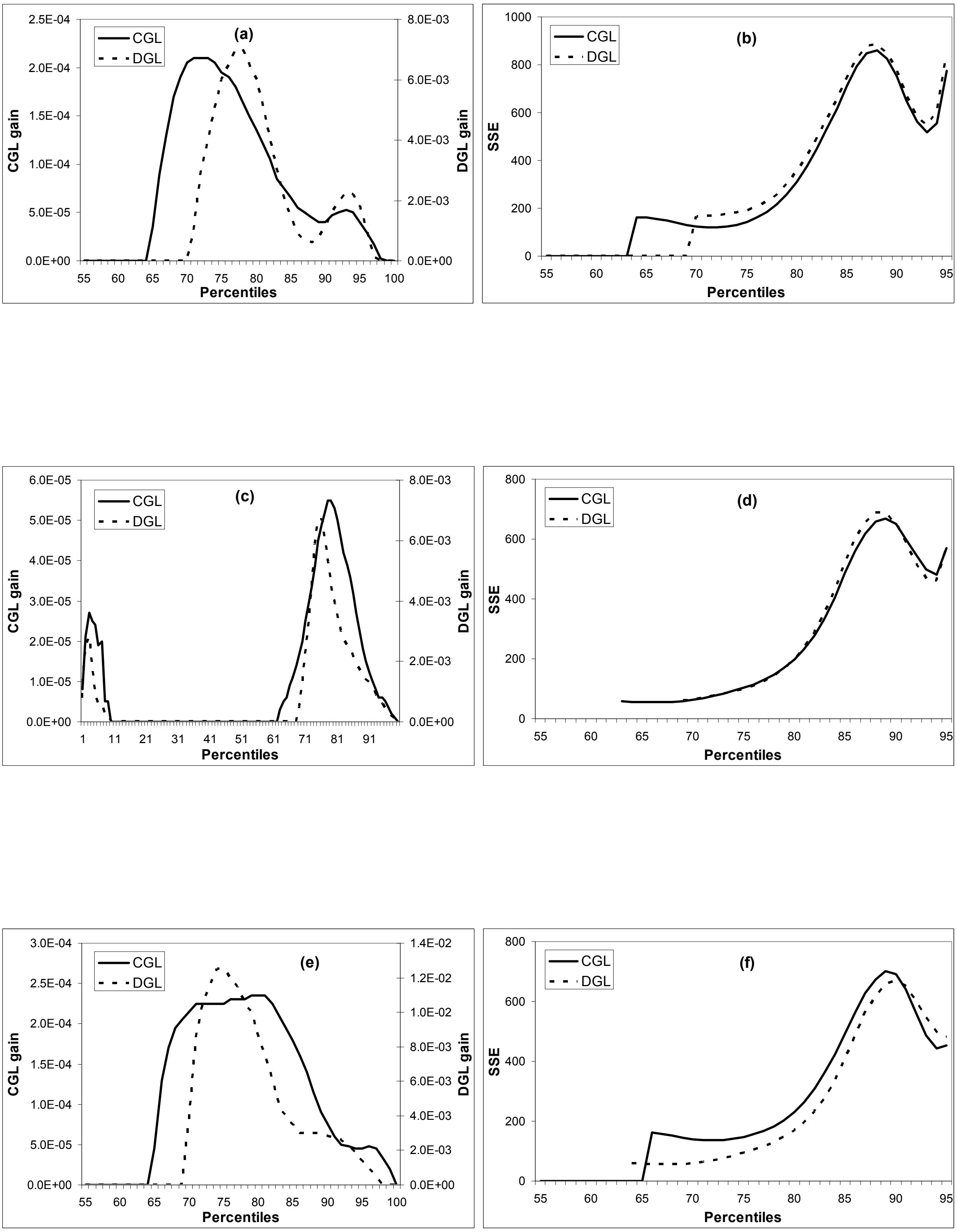
Figure 2: Determinants of the forecast error.



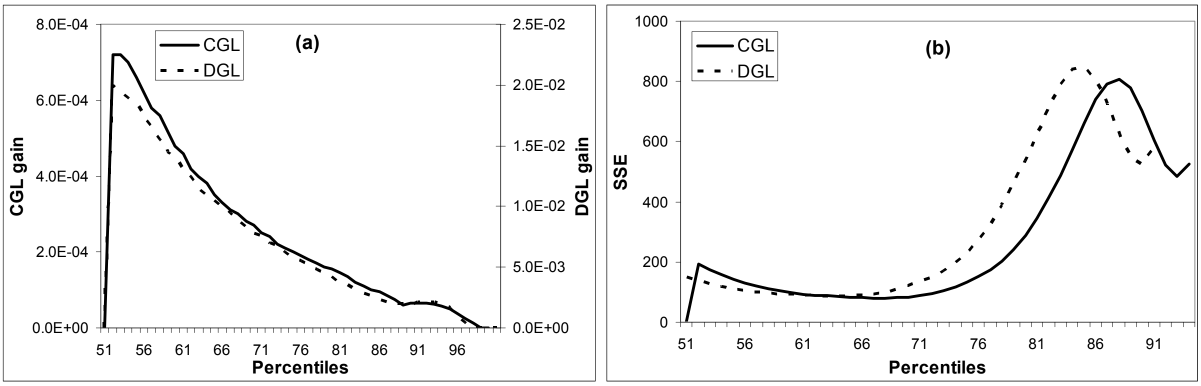
Figures 3(a)-(b): Adaptive expectations.



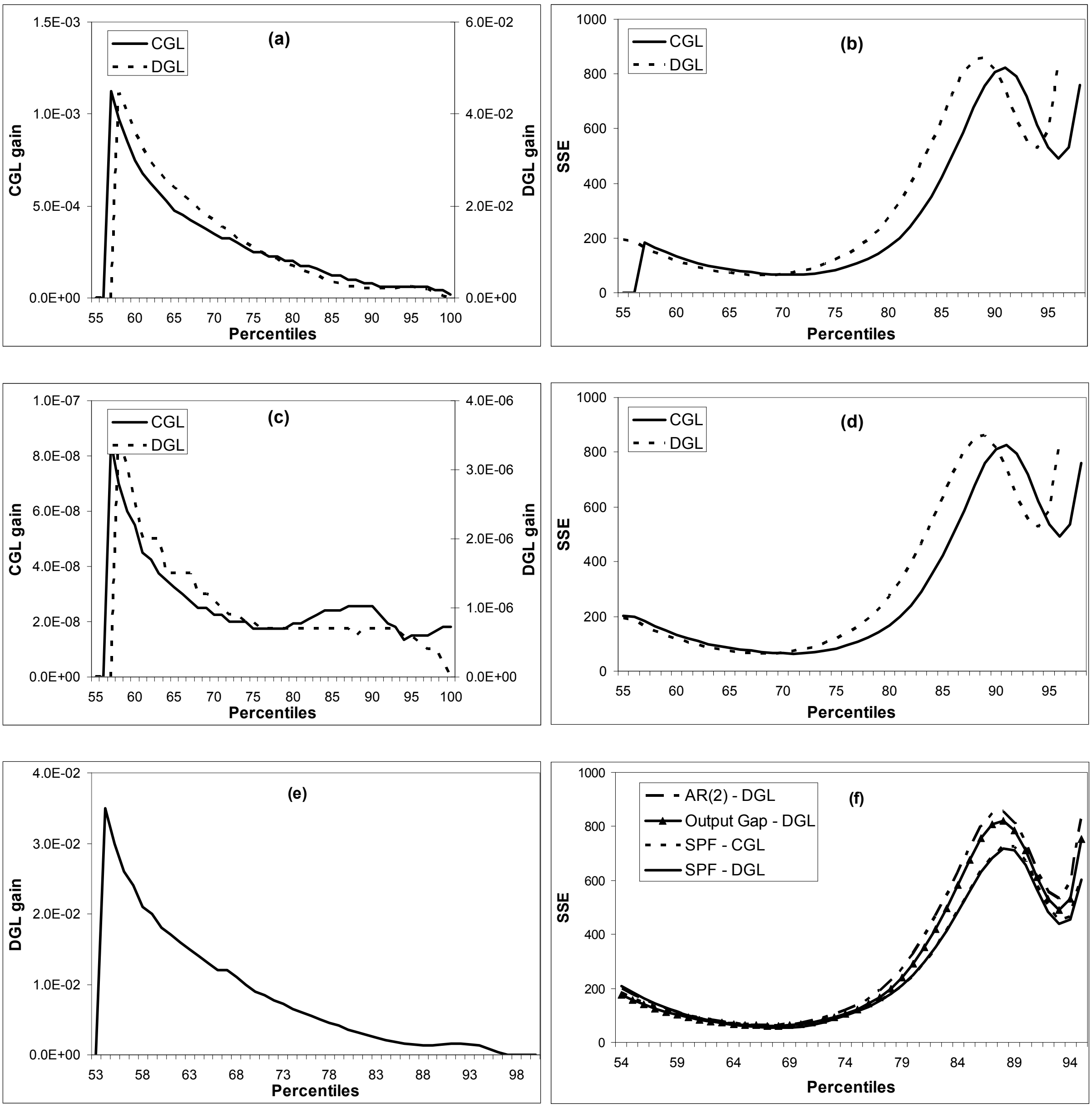
Figures 4(a)-(b): Sticky information: static and dynamic case



Figures 5(a)-(f): Adaptive learning - standard updating and iterative representation.



Figures 6(a)-(b): Adaptive learning - updating mechanism based on expected future errors.



Figures 7(a)-(f): Adaptive learning - updating mechanism based on future errors.

	Adaptive Learning			
	CGL		DGL	
	Max. Gain	Min. SSE	Max. Gain	Min. SSE
PLM with π_t and $\pi^F_{t\|t-12}$	8.00E-04	60.3	0.035*$(1/t)$	53.1
PLM with π_t and y_t	1.15E-03	61.6	0.051*$(1/t)$	61.4
PLM with π_t, π_{t-1}, π_{t-2}	4.75E-04	62.6	0.025*$(1/t)$	63.3

Table 1: Maximum gain for different PLMs (constant and decreasing gain learning).

		Forecast Range							
			LHS		Center			RHS	
Expectation Formation	*Information Set*	*1 − 10*	*11 − 30*	*31 − 49*	*50 − 55*	*56 − 66*	*67 − 72*	*73 − 90*	*91 − 100*
Rational Exp.					$\surd$				
Adaptive Exp.				*S*				*S*	*S*
Static Exp.		$\surd$	*S*						
Sticky info				*S*	*S*				*S*
Dyn. Sticky info						$\surd$	*S*		
Adapt. Learning -V1		*S*						$\surd$	*S*
Adapt. Learning -V4						*S*	$\surd$		
	Output Gap	*P*			*F*				
	Inflation		*P*	*F*	*F*	*F*	*O*	*O*	*O*
	Past Forecasts	*F*	*F*	*F*					
	SPF			*P*	*F*	*F*	*F*		
Legend: $\surd$ =best model; *S* =some evidence; *P* =partial reliance; *F* =full reliance; *O* =over-reaction									

Table 2: Heterogeneity in Inflation Expectations.

6. Appendix

Table A1: Sticky information - LSTAR model.

Perc.	Upsilon	c	Lambda	t-test	Perc.	Upsilon	c	Lambda	t-test	Perc.	Upsilon	c	Lambda	t-test
1	0.300	9.000	0.009	1.109	34	0.340	12.000	0.024	1.407	67	0.240	2.650	0.062	2.867
2	0.350	9.000	0.011	1.168	35	10.000	4.340	0.019	1.330	68	0.250	2.650	0.060	2.900
3	0.370	9.000	0.009	1.003	36	10.000	4.310	0.023	1.503	69	0.260	2.650	0.059	2.950
4	0.370	9.000	0.007	0.899	37	5.820	8.300	0.028	1.697	70	0.260	2.630	0.060	3.011
5	4.000	7.600	0.005	0.749	38	6.040	8.300	0.033	1.897	71	0.270	2.600	0.060	3.073
6	4.000	7.600	0.003	0.625	39	6.480	8.300	0.038	2.086	72	0.270	2.600	0.061	3.125
7	4.000	7.600	0.003	0.532	40	7.000	8.300	0.044	2.254	73	0.280	2.590	0.061	3.160
8	3.560	7.200	0.002	0.433	41	10.000	6.900	0.051	2.443	74	0.280	2.590	0.062	3.176
9	2.860	7.400	0.001	0.314	42	10.000	6.900	0.059	2.684	75	0.280	2.560	0.062	3.171
10	0.610	3.670	-0.001	-0.236	43	10.000	6.900	0.070	2.959	76	0.260	2.050	0.062	3.151
11	0.580	3.670	-0.001	-0.261	44	10.000	6.900	0.082	3.245	77	0.250	1.730	0.062	3.122
12	2.690	9.000	-0.002	-0.399	45	10.000	6.900	0.094	3.526	78	0.240	1.680	0.062	3.092
13	3.720	9.200	-0.003	-0.634	46	10.000	6.900	0.108	3.792	79	1.740	6.500	0.049	3.101
14	3.560	9.200	-0.003	-0.620	47	10.000	6.900	0.121	4.039	80	1.770	6.500	0.051	3.119
15	3.640	9.200	-0.003	-0.606	48	10.000	6.900	0.133	4.253	81	1.740	6.500	0.052	3.137
16	3.660	9.200	-0.003	-0.614	49	10.000	6.900	0.142	4.412	82	1.590	6.500	0.054	3.158
17	3.640	9.200	-0.003	-0.612	50	3.690	7.400	0.151	4.509	83	0.800	7.200	0.060	3.200
18	3.640	9.200	-0.003	-0.560	51	3.180	7.400	0.154	4.539	84	0.660	7.200	0.064	3.261
19	3.680	9.200	-0.002	-0.444	52	2.020	7.500	0.154	4.498	85	0.580	7.200	0.068	3.310
20	2.960	10.500	0.002	0.332	53	1.460	7.500	0.151	4.398	86	0.540	7.200	0.068	3.281
21	2.680	10.500	0.002	0.390	54	1.230	7.500	0.143	4.243	87	0.530	7.100	0.063	3.157
22	2.760	10.700	0.003	0.461	55	1.080	7.500	0.133	4.051	88	0.530	6.800	0.057	2.984
23	3.300	10.700	0.004	0.588	56	0.960	7.400	0.121	3.848	89	0.510	6.600	0.049	2.766
24	4.180	10.700	0.005	0.677	57	0.850	7.400	0.110	3.651	90	0.480	6.500	0.041	2.512
25	5.140	10.700	0.007	0.783	58	0.180	1.730	0.131	3.481	91	0.470	6.500	0.032	2.217
26	5.720	10.700	0.008	0.858	59	0.190	2.050	0.120	3.353	92	0.480	6.500	0.024	1.957
27	6.120	10.700	0.009	0.896	60	0.200	2.360	0.109	3.241	93	0.520	6.500	0.018	1.734
28	6.400	10.700	0.010	0.990	61	0.210	2.560	0.099	3.138	94	0.780	6.500	0.014	1.576
29	10.000	10.700	0.012	1.089	62	0.210	2.580	0.090	3.048	95	1.290	6.500	0.013	1.558
30	10.000	10.700	0.014	1.161	63	0.220	2.590	0.082	2.969	96	8.000	2.810	0.014	1.626
31	10.000	10.700	0.016	1.260	64	0.220	2.590	0.075	2.906	97	8.000	2.820	0.015	1.675
32	10.000	10.700	0.017	1.283	65	0.230	2.630	0.069	2.868	98	8.000	2.820	0.014	1.624
33	0.300	12.000	0.021	1.277	66	0.230	2.650	0.065	2.856	99	8.000	2.820	0.010	1.444

Table A2: Percentile Time Series Regression

Percentile	α	AR(1)	Hor. Spread	ΔCycle	SPF Forcast Err.	ΔInflation	Adj R^2	DW	LM
5	0.725	0.831	-0.212	-0.012	0.411	0.569	0.913	0.877	96.006
	5.767	31.100	-9.282	-0.350	7.493	9.843			
	0.000	0.739	0.001	0.001	0.161	0.012			
20	0.377	0.882	-0.110	0.039	0.292	0.545	0.878	0.484	177.147
	3.598	28.692	-5.047	1.232	6.058	10.199			
	0.000	0.749	-0.005	-0.004	0.131	0.008			
35	0.536	0.714	-0.130	0.055	0.235	0.530	0.737	0.662	141.077
	5.703	15.311	-4.847	1.431	3.631	7.897			
	0.000	0.484	-0.005	-0.007	0.148	0.121			
50	0.098	0.213	-0.034	0.060	0.493	0.174	0.620	0.526	168.881
	1.984	3.634	-1.333	1.652	6.884	2.503			
	0.000	0.099	-0.004	-0.014	0.449	0.097			
65	-0.888	0.219	-0.006	0.056	0.254	0.428	0.751	0.534	167.783
	-14.176	5.284	-0.327	2.103	5.494	10.620			
	0.000	0.070	-0.001	-0.019	0.268	0.437			
80	-1.958	0.236	0.000	0.011	0.057	0.815	0.703	0.884	100.847
	-16.617	6.523	-0.003	0.252	1.076	17.487			
	0.000	0.047	0.000	-0.003	0.033	0.630			
95	-7.060	0.326	0.047	0.108	-0.321	1.240	0.619	1.112	67.128
	-16.674	8.625	1.007	1.583	-3.972	17.933			
	0.000	0.115	0.005	-0.011	-0.087	0.603			

First row: coefficient; Second row: t-test; Third row: partial contribution to R^2

Figure A1: Coefficients for Percentile Time Series Regression