



**HAL**  
open science

# Heterogeneity, Learning and Information Stickiness in Inflation Expectations

Damjan Pfajfar, Emiliano Santoro

► **To cite this version:**

Damjan Pfajfar, Emiliano Santoro. Heterogeneity, Learning and Information Stickiness in Inflation Expectations. *Journal of Economic Behavior and Organization*, 2010, 75 (3), pp.426. 10.1016/j.jebo.2010.05.012 . hal-00849412

**HAL Id: hal-00849412**

**<https://hal.science/hal-00849412>**

Submitted on 31 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Accepted Manuscript

Title: Heterogeneity, Learning and Information Stickiness in Inflation Expectations

Authors: Damjan Pfajfar, Emiliano Santoro

PII: S0167-2681(10)00094-6  
DOI: doi:10.1016/j.jebo.2010.05.012  
Reference: JEBO 2547

To appear in: *Journal of Economic Behavior & Organization*

Received date: 23-1-2009  
Revised date: 16-4-2010  
Accepted date: 11-5-2010

Please cite this article as: Pfajfar, D., Santoro, E., Heterogeneity, Learning and Information Stickiness in Inflation Expectations, *Journal of Economic Behavior and Organization* (2008), doi:10.1016/j.jebo.2010.05.012

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Heterogeneity, Learning and Information Stickiness in Inflation Expectations<sup>☆</sup>

Damjan Pfajfar\*

*CentER, EBC, University of Tilburg, Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands.*

Emiliano Santoro

*Department of Economics and EPRU, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen, Denmark.*

*CIFREM, University of Trento, Via Rosmini 70, I-38122 Trento, Italy.*

---

## 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. To account for the evolution of the cross-section of inflation forecasts over time and 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

---

<sup>☆</sup>We would like to thank the Editor, two anonymous referees, Sean Holly, Seppo Honkapohja, Hashem Pesaran, Cars Hommes, Tiago Cavalcanti, Domenico Delli Gatti, Chryssi Giannitsarou, 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 and invaluable discussion. The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7/2007-2013) under Socio-economic Sciences and Humanities, grant agreement N. 225408. All remaining errors are ours. The paper was previously circulated under the title "Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment".

\*Corresponding author

*Email addresses:* [D.Pfajfar@uvt.nl](mailto:D.Pfajfar@uvt.nl) (Damjan Pfajfar), [emiliano.santoro@econ.ku.dk](mailto:emiliano.santoro@econ.ku.dk) (Emiliano Santoro)

*URL:* <http://center.uvt.nl/staff/pfajfar/> (Damjan Pfajfar),  
<http://web.econ.ku.dk/esantoro/> (Emiliano Santoro)

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:* Adaptive Learning, Sticky Information, Survey Expectations

---

Accepted Manuscript

## 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 whereby consumers update their inflation expectations from the media, which are assumed to transmit the expectations of the professional forecasters. Mankiw and Reis (2002) suggest 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 expect-

tations,<sup>1</sup> 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 different agent-types over the cross-sectional spectrum of responses in 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. We estimate the sticky information model put forward by Carroll (2003a, 2003b) for each agent-type. Only less than 10% of the forecasts reflect regular information updating. We then augment the epidemiological framework 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 when inflation "matters". We show that this mechanism fits rather well the forecasts in the upper end of the distribution, which reflect greater attentiveness 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. The first subsample is used to estimate the initial values of the parameters in the Perceived Law of Motion (PLM). In turn, these values represent the starting point for the recursive estimation of the gain parameter in the second part of the sample. The main practical inconvenience of this approach is that it does not allow the researcher to fully exploit the data available. In addition, this procedure still bears the risk that learning dynamics could result just as a statistical artifact due to a non-optimal initialization. Our approach abstracts from this criticism, as we search for the combination of initial values and gain parameter that provide the closest fit of the empirical

---

<sup>1</sup>See, for a review of these tests, Pesaran (1987), Bakhshi and Yates (1998) and Mankiw et al. (2004).

density, thus preserving the sample structure and optimizing the initialization.<sup>2</sup> Our results suggest that consumers' expectations on the right hand side (RHS) of the median forecast display adaptive learning, whereas forecasts on the left hand side (LHS) of the median 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 those 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 the forecasts on the LHS of the median forecast are highly static and entail systematic forecast 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 statistics considered as relevant for the forecasting process. Conversely, forecasts on the RHS tend to over-react to changes in the current rate of 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, expectations on the RHS of the median forecast 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,<sup>3</sup> (ii) relying on different information sets or (iii) entailing different capacities to process information. Some theoretical studies have introduced heterogeneous expectations in standard

---

<sup>2</sup>Orphanides and Williams (2003, 2005a, 2005b) and Milani (2007) have provided some empirical support for adaptive learning dynamics in DSGE models.

<sup>3</sup>Namely, agents could have different underlying assumptions about the structure of the economy or different parameterizations (or priors) of the same model.

macroeconomic models, such as in the New Keynesian framework (Branch and McGough, 2009). 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. However, most of these studies only focus on measures of central tendency to assess the degree of heterogeneity in private agents' forecasts. We show that this approach entails some fundamental fallacies if the distribution of forecasts is not time invariant and displays substantial asymmetry and multimodality. 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 set 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 information stickiness; Section 4 summarizes and discusses the main empirical results; the 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.<sup>4</sup> 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 expected changes in the price level: (i) firstly, households are asked

---

<sup>4</sup>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.



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 statement about the expected change.<sup>5</sup>

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 intervals of expected price changes, depending on their relative size with respect to the overall sample size. Only a negligible proportion of "do not know" responses is excluded.

As agents report unbounded inflation forecasts, we need to determine a point at both ends of the distribution beyond which observations should be excluded.<sup>6</sup> Curtin (1996) suggests two alternative truncations, namely at  $-10\%$  and  $+50\%$  and at  $-5\%$  and  $+30\%$ .<sup>7</sup> The analysis 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 1978.01-2005.02. Within this period, we explore the dynamic pattern of the empirical moments of the Michigan Survey Households' Expectations distribution (MSHE distribution hereafter). To account for the occurrence of a structural break, we pursue a parallel investigation on two subsamples, pre- and post-1988.12. This

---

<sup>5</sup>In case a respondent expects prices to stay the same, the interviewer must make sure that she does not actually have in mind a change in the price level which is analogous to that measured at the time the interview takes place.

<sup>6</sup>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 consider that the upper tail of the distribution is not only long, but also sparse and with large gaps between observations. Further technical considerations on the cut-off procedure are outlined in Curtin (1996).

<sup>7</sup>Curtin (1996) also suggests that no compelling evidence supports the choice of one truncation rule over the other.

choice allows us to take into adequate account the period of high inflation characterizing the first part of the sample and the subsequent disinflation.

### *2.1. Descriptive Statistics*

In this section the time series retrieved from the distribution of households' forecasts are plotted at the realized date and not at the time the forecast has been produced. Figure 1(a) plots the 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, consumers fail to account for the marked rise in the price level during the first Gulf War, by reacting with a consistent delay.

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

Figure 1(d) displays the 25<sup>th</sup>, the 50<sup>th</sup> and the 75<sup>th</sup> percentiles. Marked differences can be observed in the degree of volatility of different parts of the distribution. The 75<sup>th</sup> percentile appears to be remarkably stable after 1988, compared to the other percentiles. 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, thus reflecting that the threat of deflation was somewhat credible at that point in time.

Insert Figure 1 about here

Figure 1(e) reports the mean of the MSHE distribution and the mean forecast from the Survey of Professional Forecasters (SPF) 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 produces more accurate predictions, on average.

Figures 1(f)-(g) plot the higher moments of the distribution against two different cycle indicators.<sup>8</sup> The variance displays a marked counter-cyclical behavior, while skewness and kurtosis are pro-cyclical. Moreover, higher moments display more 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 dynamics of the skewness is qualitatively similar to the one of kurtosis. This evidence reflects the existence of a long and sparse positive tail characterized by strong variability.<sup>9</sup>

### 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. Our percentile time series analysis is not only motivated by its suitability to account for the asymmetry 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 period provides us with time series of equidistant statistics that describe the distribution under both a dynamic and a longitudinal perspective. We regard the expected change in the price level between time  $t - 12$  and time  $t$  as a random variable  $(\pi_{t|t-12})$  with distribution  $F(\cdot)$ . The  $k^{th}$  percentile  $(\pi_{t|t-12}^k)$  is the value below which  $k\%$  of the responses lie.<sup>10</sup> Therefore, we retrieve a set of ordered statistics for each month, i.e. 99 time series of percentiles. Given that percentile forecasts are not

---

<sup>8</sup>These consist of a HP detrended industrial production index (IPI) and an interpolated estimate from the Kuttner (1994) model of multivariate Kalman filtering.

<sup>9</sup>A further time series analysis on the cyclical properties of the higher moments of the MSHE distribution confirms our visual impression.

<sup>10</sup>Thus,  $F(\pi_{t|t-12}^k) = k$ .

produced by the same households over time, we cannot refer to them as coming from the same economic agents, as this would imply that forecasters are not allowed to switch across percentiles. For this reason, and for the sake of confining our analysis to the evolution of the distribution over time, we refer to percentile forecasts as coming from the same agent-types rather than from the same agents.<sup>11</sup>

A number of studies in the past have employed the mean or median forecasts from the Michigan Survey.<sup>12</sup> Implicitly, one motivation to focus on measures of central tendency is to remove any idiosyncratic component in the cross-section of forecasts. This principle applies 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 asymmetry and multimodality.

### 3.1. Rationality Tests

The rational expectations hypothesis (REH) can be interestingly tested with survey data on expectations<sup>13</sup> to allow for different degrees of forecast efficiency across the distribution of responses. To satisfy the REH, the forecasting procedure should not yield predictable (systematic) errors. A test of bias can be implemented by regressing the forecast error on a constant.<sup>14</sup> This allows us to check whether inflation expectations are centred around the right value:

$$\pi_t - \pi_{t|t-12}^k = \alpha + \varepsilon_t, \quad (1)$$

---

<sup>11</sup>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 studies (e.g., Curtin, 2005 and Pfajfar and Santoro, 2008) 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 state 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.

<sup>12</sup>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 et al. (2004).

<sup>13</sup>See Pesaran (1987), Mankiw et al. (2004) and Bakhshi and Yates (1998) for a review of these tests.

<sup>14</sup>See, for an application, Jonung and Laidler (1988) and Mankiw et al. (2004).

where  $\pi_t$  denotes inflation at time  $t$  and  $\pi_{t|t-12}^k$  is the  $k^{th}$  percentile from the MSHE distribution, while  $\varepsilon_t$  is a serially uncorrelated error term. The following regression represents a second test for rationality:

$$\pi_t = a + b\pi_{t|t-12}^k + \varepsilon_t, \quad (2)$$

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

$$\pi_t - \pi_{t|t-12}^k = a + (b - 1)\pi_{t|t-12}^k + \varepsilon_t. \quad (3)$$

Under the null of rationality, these regressions should have no predictive power.<sup>15</sup>

*Results.* The regressions based on equation (1) suggest that only the 51<sup>st</sup> – 55<sup>th</sup> (52<sup>nd</sup> – 54<sup>th</sup>) 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 et al. (2004) test the REH in Michigan Survey data. They focus on the mean and median forecast and tend to reject the null hypothesis.<sup>16</sup> 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<sup>th</sup> and 63<sup>rd</sup> percentile (56<sup>th</sup> – 62<sup>nd</sup>) are not biased at a 5% (1%) level of significance. In the 1989 – 2005 period, forecasts in the 47<sup>th</sup> – 50<sup>th</sup> (48<sup>th</sup> – 50<sup>th</sup>) percentile range are not biased at a 5% (1%) level of significance. A higher fraction of rational forecasts can be detected in the first subsample, when inflation is higher and induces a higher level of attentiveness. The second test for rationality (3) always leads to reject the null hypothesis.

### 3.2. A Further Investigation on the Forecast Error

The results reported above suggest that a substantial fraction of the inflation forecasts from the Michigan Survey are biased. We now ask to which extent different agent-types

---

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

<sup>16</sup>Roberts (1997) concludes that both Michigan and Livingston forecasts display an intermediate degree of rationality, being nor fully rational, neither entirely adaptive.

incorporate the relevant macroeconomic statistics in their forecasts. The model in equation (4) should help at unveiling the nature of the forecast error and the degree of heterogeneity in the data at hand. The set of regressors includes, among other variables, the SPF forecast error and changes in the current rate of inflation.<sup>17</sup>

$$\begin{aligned} \pi_t - \pi_{t|t-12}^k &= \alpha + \beta (\pi_{t-13} - \pi_{t-13|t-25}^k) + \delta(\pi_t - \pi_{t|t-12}^F) + \gamma \Delta X_t + \varepsilon_t, \\ k &= 1, \dots, 99; \quad X_t = \begin{bmatrix} y_t & \pi_t & (i_t - r_t) \end{bmatrix}'. \end{aligned} \quad (4)$$

where  $\pi_{t|t-12}^k$  denotes the the  $k^{th}$  percentile of the 12 months ahead expected change in prices, while  $\pi_{t|t-12}^F$  denotes the mean of the 12 months ahead expected change in prices of the mean forecast from the Survey of Professional Forecasters. Moreover,  $y_t$ ,  $\pi_t$ ,  $i_t$  and  $r_t$  denote the output gap, inflation, the real short term interest rate (3-month t-bill coupon rate) and the long term interest rate (10-year t-bond yield), respectively. Mankiw et al. (2004) and Ball and Croushore (2003) employ models similar to (4). They regress the 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 an inefficient exploitation of information from past forecasts. 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 variability in the dependent variable (Scherrer, 1984).<sup>18</sup> Table A1 and Figure A1 in the appendix report the estimated coefficients over the cross-sectional spectrum of responses.

---

<sup>17</sup>This equation could also be considered as a test of rationality. The test would be based on the null hypothesis  $H_0 : \alpha = \beta = \gamma = 0$ . Moreover, to account for the sluggishness in the transmission of information from professional forecasts to households' expectations (see Carroll, 2003a, 2003b), we also add several lags of the SPF median forecast. These turn out to have no explanatory power.

<sup>18</sup>The coefficient of multiple determination measures the proportion of the variance of the dependent variable  $y$  explained by the 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  $y$ .

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<sup>th</sup> percentile, declining thereafter and then assuming a U-shaped pattern, with a minimum occurring around the 55<sup>th</sup> percentile. The response to  $\pi_t - \pi_{t-12}^F$  is rather constant on the RHS of the median forecast [see Figure A1(b) in the Appendix], then hump-shaped around the 55<sup>th</sup> percentile, while it decreases in the last deciles. As to the impact of  $\Delta\pi_t$ , our estimates show that the response increases monotonically from the 51<sup>st</sup> percentile onward, thus displaying a substantial degree of over-reaction to changes in current inflation.

The total 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<sup>th</sup> percentile onward. It appears that the last observed error captures substantial variability in the LHS forecasts, which display a market degree of backward lookingness. The variability of the forecast error in this part of the cross-section is clearly not accounted for by the inclusion of changes in the current rate of inflation, while the situation is reversed as we move to the right over the cross-section of responses. In this range the variability of the forecast error is almost exclusively explained by the variance of changes in the current rate of inflation.

In the central forecast range the contribution of the lagged forecast error decreases, while the contribution of the SPF forecast error gains further importance. If we regard professional forecasters as rational agents, we can actually infer that the middle range of the empirical density is the least biased, especially around the 50<sup>th</sup> – 55<sup>th</sup> percentile. In this region of the density the SPF forecast error is almost the only relevant variable capturing the variability in 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. Nerlove et al. (1979) were the first to model expectations as an autoregressive process and labelled them as quasi-rational expectations. The following regression, conceived as a preliminary assessment of the degree of adaptiveness in households' forecasts, is equivalent to an adaptive expectations formula:<sup>19</sup>

$$\pi_{t|t-12}^k = \pi_{t-13|t-25}^k + \zeta (\pi_{t-13} - \pi_{t-13|t-25}^k) + \varepsilon_t, \quad (5)$$

Under this rule, different agent-types revise their expectations according to the last observed forecast error. Parameter  $\zeta$  is labelled as the "error-adjustment" coefficient. It captures the speed of adjustment of current forecasts to past forecast errors. As the participants to the Michigan Survey are asked to forecast inflation over the next year, the revision will be based on the inflation forecast for time  $t - 13$  (carried out at time  $t - 25$ ), corrected by the last observed forecast error,  $\pi_{t-13} - \pi_{t-13|t-25}^k$ . A word of caution is in order at this stage. In the adaptive learning approach, which will be 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. Conversely, adaptive expectations postulate that agents revise their expectations based on the past forecast error, but the error-adjustment coefficient is assumed to be constant.

### 3.3.1. Results

In Figures 3(a)-(b) we plot, for each percentile, the error-adjustment parameter and the corresponding  $R^2$  obtained from (5). In the remainder RHS (LHS) indicates the area over the cross-section of responses at the right (left) of the median forecast from the MSHE distribution. We adopt this convention for brevity of exposition.

Insert Figure 3 about here

---

<sup>19</sup>Unless otherwise indicated, in the remainder of the analysis it is assumed that agents are in period  $t - 12$  and they aim at forecasting inflation in period  $t$ , using the information set that includes variables up to period  $t - 13$ .



Overall, forecasts on the RHS behave (at least partly) in an adaptive manner, while past errors have little or no explanatory power for LHS forecasts.<sup>20</sup> This is suggested by the cross-sectional pattern of the error-adjustment parameter and the  $R^2$  for each percentile time series regression. Specifically, we observe a clear hump-shaped response between the 40<sup>th</sup> and 99<sup>th</sup> percentile, with a peak occurring at the 75<sup>th</sup> percentile.

### 3.4. Sticky Information

Carroll (2003a, 2003b) designs an epidemiological framework whereby households' inflation expectations are updated from news reports, which in turn are influenced by the expectations of the professional forecasters. His analysis on consumers' forecasts reports evidence of a rather slow diffusion process, with about 27% of the participants to the Michigan Survey updating their forecasts in every quarter.

#### 3.4.1. Testing for Sticky Information – Static Case

We estimate a simple regression in the vein of Carroll (2003a):

$$\pi_{t|t-12}^k = \lambda \pi_{t|t-12}^F + (1 - \lambda) \pi_{t-1|t-13}^k + \varepsilon_t. \quad (6)$$

The implicit assumption is that news about inflation spread slowly across agents, reaching only a fraction  $\lambda$  of the population in each period. These agents update their expectations from the SPF median forecast,  $\pi_{t|t-12}^F$ , while a remaining fraction  $1 - \lambda$  stick to their previous forecast,  $\pi_{t-1|t-13}^k$ .<sup>21</sup> Therefore, 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.<sup>22</sup>

---

<sup>20</sup>A negative estimate of the error-adjustment parameter 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 forecasts on LHS of the median forecast.

<sup>21</sup>Notice that (6) can in principle be re-written in a form similar to (5). However, the two rules are based on inherently different assumptions, as reflected by their timing. Unlike the adaptive expectations formula, under (6) agents do not update their expectations from the last observed forecast error: information is simply assumed to be condensed into news reports that transmit the views of professional forecasters (see Carroll, 2003a).

<sup>22</sup>It should be pointed out that this model is derived under the assumption that: (i) inflation follows a random walk process and (ii)  $\pi_{t|t-13}^k \approx \pi_{t-1|t-13}^k$  (see Döpke et al., 2008).

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. On the other hand, forecasts slightly above the median reflect systematic errors when inflation is lower and more stable, thus supporting the inattentiveness argument. We explore this issue in further detail in the next subsection.

*Results.* Figure 4(a) plots the estimated  $\lambda^{-1}$  for each percentile. This represents a measure of the average updating period. The estimation confirms the existence of static behavior in the information structure up to the 40<sup>th</sup> percentile. From this point up to the 91<sup>st</sup> percentile a U-shaped pattern emerges, with a minimum occurring at the 50<sup>th</sup> percentile. This translates into an average minimum updating period of 7 months. Carroll (2003a) finds similar results: his estimate of  $\lambda^{-1}$  is 11 months for the mean,<sup>23</sup> while Döpke et al. (2008) estimate an average updating period of 18 months for the Euro area. Mankiw et al. (2004) and Branch (2007) set  $\lambda$  at 0.1, which implies an average frequency of 10 months.

### 3.4.2. Testing for Sticky Information - State Dependent Coefficients

When inflation matters agents are expected to update their information set more frequently, so as to produce more accurate forecasts. In addition, in periods of marked macroeconomic turmoil the amount of media coverage is generally higher, and thus the cost of acquiring information is lower. We assume that a higher proportion of agent-types pay 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:<sup>24</sup>

$$\pi_{t|t-12}^k = \lambda F(\pi_{t-13}) \pi_{t|t-12}^F + [1 - \lambda F(\pi_{t-13})] \pi_{t-1|t-13}^k + \varepsilon_t, \quad (7)$$

---

<sup>23</sup>Mankiw and Reis (2002) consider a quarterly frequency and set  $\lambda = 0.25$  in their simulations (which corresponds to an average updating frequency of 12 months).

<sup>24</sup>For details about smooth-transition regression models see Granger and Teräsvirta (1993).

where  $F$  denotes the following logistic function:<sup>25</sup>

$$F(\pi_{t-13}) = \frac{1}{1 + \exp[-v(\pi_{t-13} - c)]}, \quad (8)$$

and  $v$  can be interpreted as a parameter measuring the speed of responsiveness, whereas  $c$  is a threshold coefficient. The approach consists of estimating  $\lambda$  through a least squares regression while running a grid search on  $v$  and  $c$ , so as 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 for every percentile. Since the SSE is always lower under the state dependent model, we can assert that the non-linear sticky information model outperforms its linear counterpart.<sup>26</sup> The responses between the 59<sup>th</sup> and 79<sup>th</sup> 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  $[\lambda F(\pi_{t-13})]^{-1}$  for the 52<sup>nd</sup> and 63<sup>rd</sup> percentile.<sup>27</sup> This is a time-varying estimate of the average updating period. The 63<sup>rd</sup> percentile forecast reflects evidence in line with the inattentiveness view. At the beginning of the sample the average updating period is rather low, as inflation is higher and so the opportunity cost of not updating information. The optimal coefficients in the transition function for this percentile are  $v = 0.21$  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<sup>nd</sup> percentile is quite different. This percentile reflects lower attentiveness only sporadically. The optimal coefficients in the transition function for the 52<sup>nd</sup> percentile are  $v = 3.18$  and  $c = 7.40$ . The interpretation of these coefficients is different from that proposed for the 63<sup>rd</sup> percentile, as  $v$  is higher than 1. In turn,  $c$  cannot be interpreted as a perceived inflation target. For

---

<sup>25</sup>We have also tried different forms of transition function, but the symmetric case we present outperforms the other alternatives in terms of SSE.

<sup>26</sup>The average absolute difference in the SSE is about 0.595.

<sup>27</sup>Table A2 in the Appendix reports the estimates for all the percentiles.

this agent-type the difference between the linear and non-linear model is negligible. The 50<sup>th</sup> – 58<sup>th</sup> 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 beyond the 80<sup>th</sup> percentile, although the LHS of the distribution is associated with a much higher average updating frequency.

Notice that the model in equation (7) can also be seen as an alternative specification to that proposed by Branch (2007). In this case the choice between different updating frequencies is modelled through a mechanism à la Brock and Hommes (1997). Branch shows that most of the agents update their information set every 3 – 6 months, while fewer agents update their information set on a monthly basis. Some agents update expectations every 9 months or even less frequently. Our results stand in partial contrast to Branch (2007). We show that information updating is less frequent on the LHS. Forecasts on the RHS behave in line with the inattentiveness argument and generally reflect lower updating frequency during periods of stable inflation (approximately every two years). By contrast, 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 detect the presence of learning dynamics in the data and 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 economic agents behave like econometricians, using the available information at the time of the forecast. Let us assume that a generic agent considers the following

perceived law of motion (PLM):<sup>28</sup>

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

where coefficients are assumed to be time-varying and follow a specific updating mechanism that will be briefly detailed. In our setting different agent-types estimate their PLM by exploiting the information available up to period  $t - 13$ . As new data become available, these update their estimates according to a constant gain learning (CGL) or a decreasing gain learning (DGL) rule. 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, each agent-type updates coefficients according to the following rule (see Evans, Honkapohja and Williams, 2010):

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

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

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta R_{t-1}^{-1} X_{t-25}' \left( \pi_{t-12} - X_{t-25} \widehat{\phi}_{t-13} \right); \quad (11)$$

$$R_t = R_{t-1} + \vartheta \left( X_{t-25} X_{t-25}' - R_{t-1} \right). \quad (12)$$

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

### 3.5.1. A Standard Updating Mechanism

To test for learning in the MSHE distribution we start by specifying the following PLM:

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

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

---

<sup>28</sup>When testing for adaptive learning it is assumed that each agent-type produces her forecast for period  $t$  at time  $t - 12$ , on the basis of the information set that includes variables up to time  $t - 13$ . Therefore, the current realization of the rate of inflation ( $\pi_{t-12}$ ) has not yet been observed when the forecast is being carried out. After observing  $\pi_{t-12}$  each agent-type proceeds to update her coefficient values.

as possible. The drawback implicit in this approach is that the initial values of  $\hat{\phi}_t$  for 12 periods have to be assumed. In the recursive estimation of the gain parameter setting the initial values for the parameters in the PLM represents the main problem: this issue 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 used to estimate the set of initial values for the parameters in the PLM. These values are then employed for the recursive estimation of the gain parameter in the second part of the sample. This approach has a main drawback, as it does not allow the researcher to fully exploit the data available. Moreover, following this procedure entails a potential risk that learning dynamics could indeed result as a statistical artifact due to a non-optimal initialization.

Our approach is not subject to this criticism, as we search for the optimal combination of initial values and gain parameter, thus preserving the sample structure and optimizing the initialization. In practice, several forecast series ( $\pi_{t|t-12}^s$ ) are simulated by means of a multidimensional grid search, under different combinations of  $\vartheta$  and  $\hat{\phi}$ . We then select the gain parameter  $\vartheta$  (or  $\iota$  under DGL) and the set of initial values  $\hat{\phi}$  that minimize the sum of squared errors (SSE), i.e.  $\left(\pi_{t|t-12}^s - \pi_{t|t-12}^k\right)^2$ .<sup>29</sup> This strategy can also be regarded as a test for learning dynamics. Importantly, if the gain is found to be positive under this method of initialization, then the series would exhibit learning for every other initialization method with a higher (or equal) gain. For this reason, in the remainder of the analysis we pose strong emphasis on the gain parameter to detect learning dynamics.<sup>30</sup>

---

<sup>29</sup>However, this approach has an obvious practical inconvenience, as running a grid search on several variables is computationally very intensive.

<sup>30</sup>A more practical consideration imposes us to look at the gain, rather than at the time-varying coefficients in the PLM, to make inference about the presence of adaptive learning in the data at hand. Considering the gains over the cross-sectional spectrum of responses allows us to establish an immediate comparison among agent-types. By contrast, comparing the time-varying parameters in the PLM for each agent-type would not be informative, as this would require to observe the estimated coefficients both over the time dimension and the longitudinal one.

*Results.* The 65<sup>th</sup>-98<sup>th</sup> 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  $2.1 \times 10^{-4}$ . This maximum is located between the 71<sup>st</sup> and 73<sup>rd</sup> percentile. The DG version of gradient learning turns out to be significant for the 70<sup>th</sup> – 96<sup>th</sup> percentile range. Within this interval the estimated gain displays properties similar to those detected under CGL. Both CGL and DGL exhibit a second minor hump on the RHS. This is more pronounced under DGL [see Figure 5(a)]. The highest gain is estimated around the 76<sup>th</sup> – 77<sup>th</sup> percentile ( $0.007t^{-1}$ ). To compare both versions of gradient learning, we plot their SSEs in Figure 5(b).<sup>31</sup> Our results suggest that the CG version of gradient learning generally provides a better fit of the forecasting process of different agent-types, especially those around the 70<sup>th</sup> 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 the agent-type relies on 12.5 years of data to forecast inflation. Pfajfar and Žakelj (2009) estimate an average gain of 0.04 with experimental data obtained within a New Keynesian model calibrated at a quarterly frequency. As in this study we explore monthly data, an estimate of  $2.1 \times 10^{-4}$  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) have recently delivered evidence suggesting that the gain parameter might actually lie in the interval [0.0015, 0.0029], a range of values much closer to our estimates.

Insert Figure 5 about here

In Figure 6 we report the time-varying coefficients in (13) for the 72<sup>nd</sup> percentile, which is representative of an agent-type whose forecasting process is consistent with adaptive learning, as shown in the analysis of the gain parameter.

---

<sup>31</sup>Results and conclusions are the same, irrespective of whether we draw our inference from the  $R^2$ , the root mean square error (RMSE), or the SSE, as these statistics are all monotonic transformations of each other.

Insert Figure 6 about here

This graph confirms that learning dynamics does have a relevance in the data at hand, as a substantial change in the parameters of the PLM takes place over the time window considered. Both parameters are positive: as to the intercept, this is rather stable in the first part of the sample, and then starts decreasing when the disinflationary policy is undertaken. The coefficient  $\hat{\phi}_{1,t-1}$  displays a minor hump during the period of high inflation, decreasing thereafter along a negative trend. Both coefficients follow a similar pattern after the change in the monetary policy regime has occurred.<sup>32</sup> Although the gain is rather small, it implies substantial changes in the coefficient value over the time window considered.

We also consider alternative PLMs. In the following formulation we replace  $\pi_{t-1}$  with  $\pi_{t-1|t-13}^k$ :

$$\pi_{t|t-12}^s = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1|t-13}^k + \varepsilon_t. \quad (14)$$

This formulation is found to provide a better fit of the data compared to (13) [see Figure 5(d)]. In this case some learning dynamics can also be detected on the LHS. The  $1^{st} - 9^{th}$  and the  $63^{rd} - 99^{th}$  percentile intervals 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. Under CGL the response pattern on the RHS is bell-shaped, with the highest gain occurring at the  $78^{th} - 79^{th}$  percentile ( $5.5 \times 10^{-5}$ ). The response under DG is also hump-shaped on the RHS, reaching the highest gain at the  $75^{th} - 76^{th}$  percentile ( $0.0067t^{-1}$ ).

### 3.5.2. An Iterative Representation of the PLM

We now introduce a PLM featuring the last observed rate of inflation as a regressor:

$$\pi_{t-12|t-13}^s = \phi_{0,t-13} + \phi_{1,t-13}\pi_{t-13} + \varepsilon_t. \quad (15)$$

We implement the following gradient learning updating algorithm:

$$\hat{\phi}_{t-12} = \hat{\phi}_{t-13} + \vartheta X'_{t-13} \left( \pi_{t-12} - X_{t-13} \hat{\phi}_{t-13} \right). \quad (16)$$

---

<sup>32</sup>In principle, the pattern of both coefficients is consistent with convergence to some value, although this is not observed over the sample we consider.



As we look at 12 months-ahead forecasts, the following rule is considered:

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

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

*Results.* This recursive algorithm delivers results similar to those obtained under (9). However, this learning method provides a slightly less accurate explanation of households' inflation forecasts. Our estimates suggest that the 65<sup>th</sup>-99<sup>th</sup> percentile range displays CGL dynamics. In this case, the optimal gains exhibit a marked hump-shaped pattern. The peak occurs at  $2.35 \times 10^{-4}$ , between the 79<sup>th</sup> and 82<sup>nd</sup> percentile [see Figure 5(e)]. In the DG case, the gain peaks at  $0.0125t^{-1}$ , between the 74<sup>th</sup> and 75<sup>th</sup> percentile. As under the previous PLM, CGL always outperforms DGL [see Figure 5(f)].

Notice that the evidence produced so far under stochastic gradient learning is confirmed after taking into account the matrix of second moments.<sup>33</sup>

### 3.5.3. An Updating Mechanism Based on Expected Future Errors

In the next updating mechanism we allow for a higher degree of forward lookingness compared to that traditionally assumed in the adaptive learning literature. We introduce a novel mechanism of expectation formation which assumes that coefficient estimates are updated with respect to new information about future inflation. New information is proxied by the mean forecast from the SPF,  $\pi_{t|t-12}^F$ . Implicitly, this model states that each agent-type updates her 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 represents a combination of adaptive learning and sticky information. We assume a PLM of the following form:

$$\pi_{t|t-12}^s = \phi_{0,t-1} + \phi_{1,t-1} \pi_{t-13} + \varepsilon_t. \quad (18)$$

---

<sup>33</sup>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 the covariance terms are quite small.

The following gradient learning updating algorithm is considered:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1} \left( \pi_{t|t-12}^F - X_{t-13} \widehat{\phi}_{t-1} \right). \quad (19)$$

*Results.* Our results suggest that learning dynamics is displayed above the 52<sup>nd</sup> percentile under CG and the 51<sup>st</sup> 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 7(b), CGL significantly outperforms DGL beyond the 65<sup>th</sup> percentile.

Insert Figure 7 about here

The highest gain is reached slightly above the median and declines thereafter. Compared to the previous updating algorithms, a higher proportion of forecasts are consistent with (19). This signals that, despite the fact that some evidence of learning dynamics can be detected, substantial forward lookingness characterizes the updating procedure. Moreover, as it will emerge in the implementation of the next updating rule, the distribution of forecasts reflects a higher informational content, compared to that proxied by the views of the professional forecasters.

#### 3.5.4. An Updating Mechanism Based on Future Errors

We now allow for the possibility that more (and better) information about future inflation developments can be accessed than that reflected in the SPF forecasts. One motivation for this concern is that several studies have documented the presence of herding behavior in the predictions of the professional forecasters.<sup>34</sup> Under these circumstances their predictions might not reflect the information available in the relevant macroeconomic statistics. We proxy this information with  $\pi_t$ . We consider the PLM in Equation (18). We first implement a gradient learning updating algorithm:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1} \left( \pi_t - X_{t-13} \widehat{\phi}_{t-1} \right). \quad (20)$$

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

---

<sup>34</sup>Scharfstein and Stein (1990), Banerjee (1992) and Zwiebel (1995) argue that forecasters are occasionally afraid to deviate from the majority or the consensus opinion. Pons-Novell (2003) provide empirical evidence on this.

Insert Table 1 and Figure 8 about here

*Results.* The forward-looking updating mechanism (20) allows us to assess the importance of learning from new information, compared to previous rules characterized by a backward-looking perspective. Our results suggest that forecasts on the RHS can be associated with this version of adaptive learning. To support this evidence, we also explore a wider set of potential PLMs. We start with (13). In this case the data display learning dynamics from the 55<sup>th</sup> (56<sup>th</sup>) percentile onward under CG (DG). In both cases, the gain immediately jumps to the highest value and decreases over the cross-section thereafter. The highest gain is estimated at  $1.125 * 10^{-3}$  and the lowest SSE is reached at the 68<sup>th</sup> percentile [see Figures 8(a)-(b)]. Compared to the estimates obtained under the first updating scheme, this gain can be regarded as a more realistic one, 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, if we exclude those in the 63<sup>rd</sup>-69<sup>th</sup> interval.

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 [see Figures 8(c)-(d)]. 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 the mean forecast from the SPF. We find that the PLM implementing the SPF inflation forecasts outperforms the other options, especially under DGL [see Figure 8(e)]. The pattern of the optimal gain is quite similar across competing PLMs. The general impression is that agent-types between the 54<sup>th</sup> and 98<sup>th</sup> percentile behave in accordance with an adaptive learning scheme based on the SPF inflation forecasts. Figure 8(f) plots the SSEs, whereas Table 1 reports the maximum gains. The optimal gain under CGL is estimated between 0 and 0.051. Moreover, we obtain a better fit of the data under DGL. Overall, these results confirm that forecasts on the RHS are better explained by this learning mechanism, compared to "more traditional" learning schemes.

#### 4. Discussion

Our analysis highlights the presence of a marked degree of heterogeneity in the process of expectation formation. We identify three regions over the cross-section of the MSHE density that correspond to different underlying mechanisms of expectation formation. On the one hand, we can consider the forecast range on the LHS as that characterized by agent-types that do not exploit the relevant information when forecasting inflation. On the other hand, predictions on the RHS reflect significant over-reaction to the flow of information about future inflation. Intuitively, forecasts in the middle range of the density are nearly unbiased. Table 2 reports, for each percentile range, 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

Forecasts of the LHS 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<sup>th</sup> 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 an adaptive learning scheme whereby parameters are updated with respect to past errors. To conclude, forecasts in this sub-interval are highly inertial 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<sup>th</sup> and 30<sup>th</sup> percentile. Households' forecasts in this interval 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 supported by the data. We could characterize this sub-interval with a PLM featuring an intercept, past forecasts and past inflation. The third sub-interval incorporates forecasts in the 31<sup>th</sup> – 49<sup>th</sup> percentile range. Information updating occurs on a more regular basis, especially after the 40<sup>th</sup> percentile. Past inflation is fully exploited into the PLM, as well as information on future developments in inflation. 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 an intercept, last period's forecast and inflation, and the mean forecast from the SPF.

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

The RHS 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 in this range display inattentiveness. The average updating frequency ranges between 8 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 expectations, as the forecast errors are increasingly associated with changes in the current rate of inflation as the percentile order increases. The second sub-interval on the RHS 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 such that forecasts are updated with respect to the arrival of new information (proxied by  $\pi_{t|t-12}^F$ ). The forecasts between the  $73^{rd}$  and  $90^{th}$  percentile can be grouped into the third sub-interval. Their dynamics can be replicated rather accurately under a CGL algorithm with a PLM featuring past forecasts. The 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 the current rate of inflation acquire a predominant role in the determination of the forecast errors in this area. The last range of forecasts on the RHS can be placed above the  $91^{st}$  percentile. This can be associated with a DG version of learning, whereby coefficients are updated with respect to new information. However, expectations are updated rather infrequently and the forecast errors are almost exclusively explained by changes in the current rate of inflation.

A further point relates to the connection between the evidence we report and the experimental evidence on inflation expectations. Laboratory experiments have shown that simple

forecasting models, such as linear AR(1) rules or simple trend extrapolation rules, fit the individual forecasting behavior quite well (Hey, 1994; Hommes et al., 2005). Hommes et al. (2005) consider an asset pricing framework and find that about 50% of the subjects in the experiment behave in line with a trend extrapolation rule. Also studies on survey data on exchange rates expectations, such as Frankel and Froot (1987), have shown that adaptive expectations or trend extrapolation rules fit these data quite well.<sup>35</sup> We have assessed the performance of different trend extrapolation rules in accounting for the Michigan Survey participants' forecasting behavior. Neither of the rules we employ outperform the sticky information model or the adaptive expectations formula. The discrepancy between our results and the experimental evidence is likely to be driven by two main factors. First, we argue that cycles or, more generally, clear trends are observed in experimental data (or in survey data on asset prices and exchange rates), while Michigan Survey data on inflation expectations and CPI inflation exhibit a clear trend just in the first part of the sample, when the disinflation took place. Second, it is our intuition that a trend extrapolation rule finds little empirical support if the data display volatility at a time-frequency higher than the length of the forecast horizon, as monthly changes in inflation tend to be "smoothed out" over the past year. As the participants to the Michigan Survey are asked to forecast 12 months-ahead inflation and data are available at a monthly frequency, the contribution of the last observed change in inflation ( $\pi_{t-13} - \pi_{t-25}$ ) to the forecasting process may be negligible.

## 5. 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. To account for the marked degree of asymmetry and multimodality in the empirical

---

<sup>35</sup>More recently, Pfajfar and Žakelj (2009) conduct an experimental study on agents forecasting inflation within a New Keynesian setting. They find that approximately 35 – 45% of the subjects can be regarded as rational and that around 30 – 35% of them predominantly extrapolate a trend. Moreover, 15 – 25% of the subjects mostly behave in line with adaptive learning and sticky information models, while only 5% of the forecasts reflect adaptive behavior, i.e. these forecasts are updated with respect to the last observed error.

density, we apply our techniques to the entire cross-sectional range of forecasts.

We first extend the epidemiological framework proposed by Carroll (2003a) to account for the possibility that agents are more likely to update their expectations on a more regular basis during 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 for the empirical implementation of 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 coefficients in the PLM and gain parameter, thus preserving the sample structure and optimizing the initialization procedure. This technique 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 based on their (expected) future errors. 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 allow 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" fraction of forecasts around the median, and a fraction of forecasts on the right hand side of the median which are produced in line with models of adaptive learning and sticky information.

## References

- Bakhshi, H., Yates, A., 1998. Are U.K. Inflation Expectations Rational? Technical Report 81. Bank of England Working Paper.
- Ball, L., Croushore, D., 2003. Expectations and the effects of monetary policy. *Journal of Money, Credit and Banking* 35, 473–484.

- Banerjee, A.V., 1992. A simple model of herd behavior. *The Quarterly Journal of Economics* 107, 797–817.
- Bernanke, B., 2007. Inflation Expectations and Inflation Forecasting. Technical Report. Speech at National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts.
- Branch, W.A., McGough, B., 2009. A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control* 33, 1036–1051.
- Branch, W.A., 2004. The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal* 114, 592–621.
- Branch, W.A., 2007. Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control* 31, 245–276.
- Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. *Econometrica* 65, 1059–1096.
- Carceles-Poveda, E., Giannitsarou, C., 2007. Adaptive learning in practice. *Journal of Economic Dynamics and Control* 31, 2659–2697.
- Carroll, C.D., 2003a. Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics* 118, 269–298.
- Carroll, C.D., 2003b. The Epidemiology of Macroeconomic Expectations. NBER Working Papers 8695.
- 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.
- Curtin, R., 2005. Inflation Expectations: Theoretical Models and Empirical Tests. Mimeo. University of Michigan.



- Döpke, J., Doovern, J., Fritsche, U., Slacalek, J., 2008. The dynamics of european inflation expectations. *Topics in Macroeconomics* 8, 1540–1540.
- Eusepi, S., Preston, B., 2008. Expectations, Learning and Business Cycle Fluctuations. NBER Working Paper 14181.
- Evans, G.W., Honkapohja, S., Williams, N., 2010. Generalized stochastic gradient learning. *International Economic Review* 51, 237–262.
- Evans, G.W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press.
- Fisher, I., 1930. Theory of Interest. Macmillan, New York.
- Frankel, J.A., Froot, K.A., 1987. Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review* 77, 133–53.
- Granger, C.W.J., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. Oxford University Press, New York.
- Hey, J.D., 1994. Expectations formation: Rational or adaptive or ...? *Journal of Economic Behavior & Organization* 25, 329–349.
- Hommel, C., Sonnemans, J., Tuinstra, J., van de Velden, H., 2005. Coordination of expectations in asset pricing experiments. *Review of Financial Studies* 18, 955–980.
- Jonung, L., Laidler, D., 1988. Are perceptions of inflation rational? Some evidence for Sweden. *American Economic Review* 78, 1080–87.
- Kuttner, K.N., 1994. Estimating potential output as a latent variable. *Journal of Business of Economic Statistics* 12, 361–68.
- Mankiw, N.G., Reis, R., Wolfers, J., 2004. Disagreement about inflation expectations. NBER Macroeconomics Annual 2003 18, 209–248.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics* 117, 1295–1328.

- Milani, F., 2005. Adaptive Learning and Inflation Persistence. *Macroeconomics* 0506013. EconWPA.
- Milani, F., 2007. Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics* 54, 2065–2082.
- Nerlove, M.L., Grether, D.M., Carvalho, J.L., 1979. *Analysis of Economic Time Series: A Synthesis*. New York: Academic Press, Inc. revised edition, 1995 edition.
- Orphanides, A., Williams, J.C., 2003. Imperfect Knowledge, Inflation Expectations, and Monetary Policy. NBER Working Papers 9884.
- Orphanides, A., Williams, J.C., 2005a. The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations. *Journal of Economic Dynamics and Control* 29, 1927–1950.
- Orphanides, A., Williams, J.C., 2005b. Inflation scares and forecast-based monetary policy. *Review of Economic Dynamics* 8, 498–527.
- Pesaran, M.H., 1987. *The Limits to Rational Expectations*. Basil Blackwell, Oxford. reprinted with corrections 1989 edition.
- Pfajfar, D., Santoro, E., 2008. Asymmetries in Inflation Expectation Formation Across Demographic Groups. Cambridge Working Papers in Economics 0824. Faculty of Economics, University of Cambridge.
- Pfajfar, D., Žakelj, B., 2009. Experimental Evidence on Inflation Expectation Formation. Discussion Paper 2009-07. Tilburg University, Center for Economic Research.
- Pons-Novell, J., 2003. Strategic bias, herding behaviour and economic forecasts. *Journal of Forecasting* 22, 67–77.
- Roberts, J.M., 1997. Is inflation sticky? *Journal of Monetary Economics* 39, 173–196.
- Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. *American Economic Review* 80, 465–79.

Scherrer, B., 1984. Biostatistique. Gatan Morin, Chicoutimi.

Zwiebel, J., 1995. Corporate conservatism and relative compensation. *Journal of Political Economy* 103, 1–25.

Accepted Manuscript

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

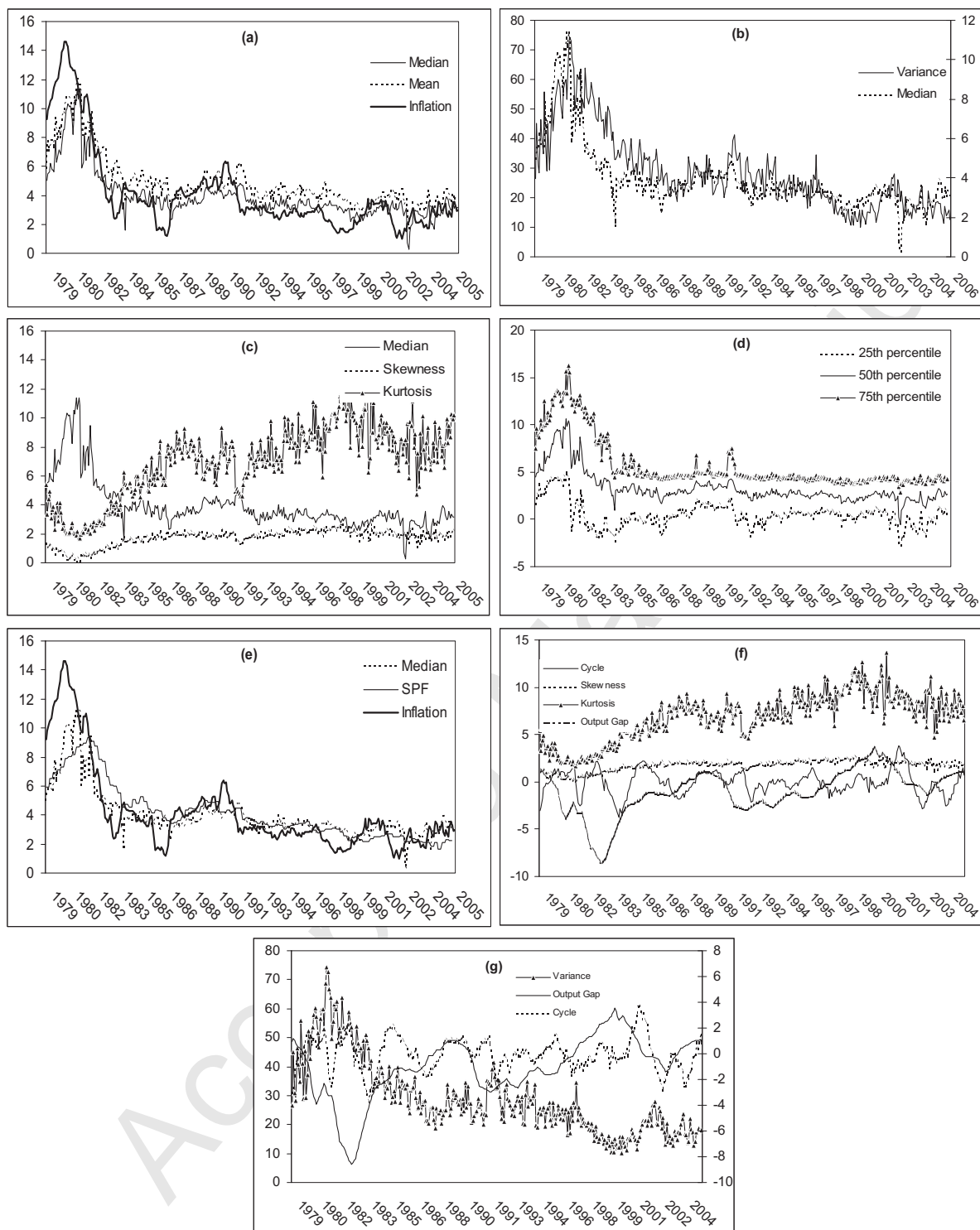
Table 1: Maximum gain for different PLMs (constant and decreasing gain learning).

	Forecast Range													
	LHS			Middle Range			RHS							
<i>Expectation Formation</i>	<i>Information Set</i>	<i>1 - 10</i>	<i>11 - 30</i>	<i>31 - 49</i>	<i>50 - 55</i>	<i>56 - 66</i>	<i>67 - 72</i>	<i>73 - 90</i>	<i>91 - 99</i>					
Rational Exp.					√									
Adaptive Exp.				S				S						
Static Exp.	√		S											
Sticky Info.			S		S									S
Dyn. Sticky Info.						√	S							
Adapt. Learning - V1 <sup>(a)</sup>		S						√						
Adapt. Learning - V4 <sup>(b)</sup>						S	√							
Output Gap		P				F								
Inflation			P	F	F	F	O	O	O					
Past Forecasts		F	F	F										
SPF				P	F	F	F							

Legend: √ = Best Fit (Lowest SEE); S = Some Evidence; P = Partial Reliance; F = Full Reliance; O = Over-reaction

Adaptive Learning: (a) Updating Algorithm in eq. (11); (b) Updating Algorithm in eq. (20).

Table 2: Heterogeneity in Inflation Expectations.



Figures 1(a)-1(g): Empirical moments of the MSHE distribution. Notes: the empirical moments of the MSHE distribution are plotted at the realized date, i.e. at the time the 12 months-ahead forecast refers to. *Cycle* and *Output Gap* in Figure 1(g) denote the HP-detrended industrial production index (IPI) and an interpolated estimate from the Kuttner (1994) model of multivariate Kalman filtering, respectively.

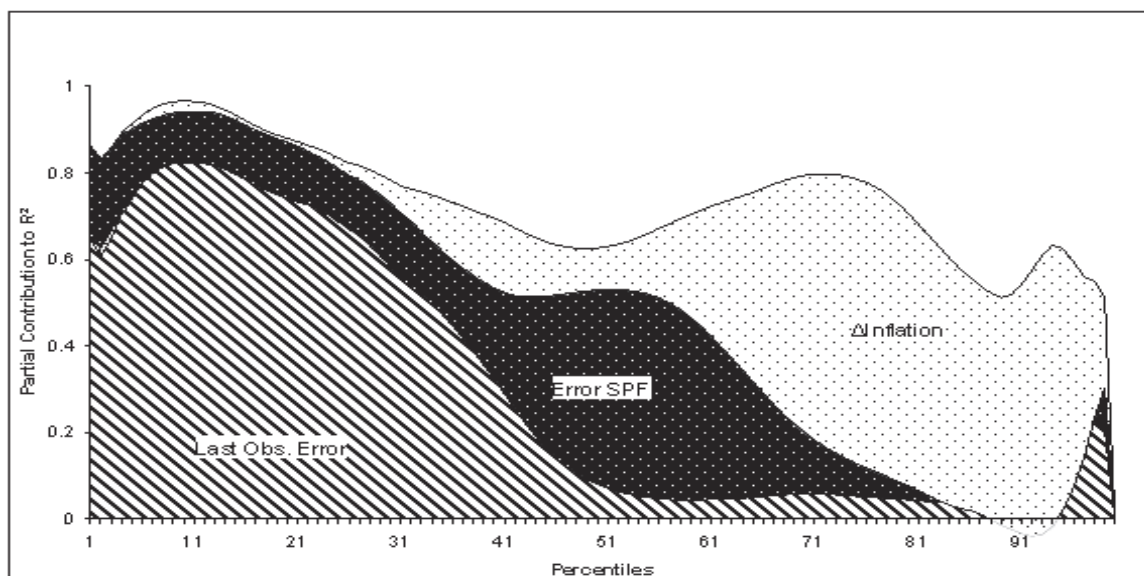


Figure 2: Partial  $R^2$  associated with the regressors included in Equation (4). Legend:  $LastObs.Error = \pi_{t-13} - \pi_{t-13|t-25}^k$ ,  $ErrorSPF = \pi_t - \pi_{t|t-12}^F$ ,  $\Delta Inflation = \Delta \pi_t$ . Notes: a least squares estimation is carried out for each percentile; the partial  $R^2$  accounts for the contribution of each regressor to the explanation of the variation in the dependent variable; all standard errors are corrected for heteroskedasticity and serial correlation using a Newey-West procedure.

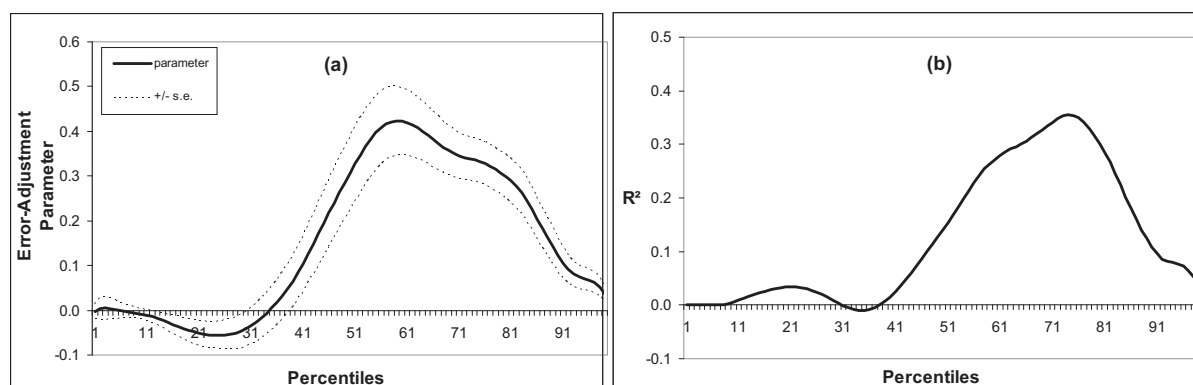


Figure 3: (a) Estimated error-adjustment parameter and (b)  $R^2$  from the estimation of Equation (5). Notes: a least squares estimation is carried out for each percentile; the dotted lines in Figure 3(a) are two standard error bands for the estimated parameter; all standard errors are corrected for heteroskedasticity and serial correlation using a Newey-West procedure.



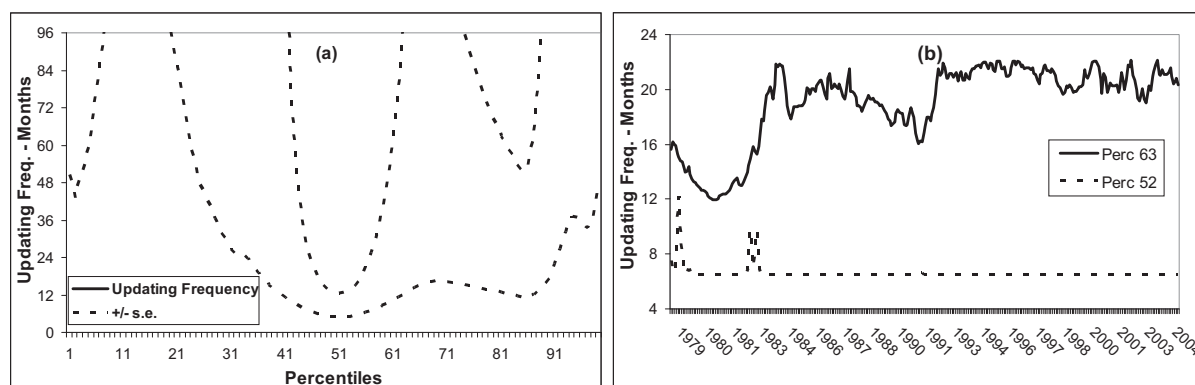


Figure 4: Estimated monthly updating frequency from the sticky information model. Notes: Figure 4(a) reports the least squares estimate of the updating frequency in the static case, while Figure 4(b) reports the time-varying frequency of information updating for the 63<sup>rd</sup> and 52<sup>nd</sup> percentile from the estimation of the state dependent framework, specified as a logistic smooth-transition autoregressive (LSTAR) model; the dotted lines in Figure 4(a) are two standard error bands for the estimated parameter.

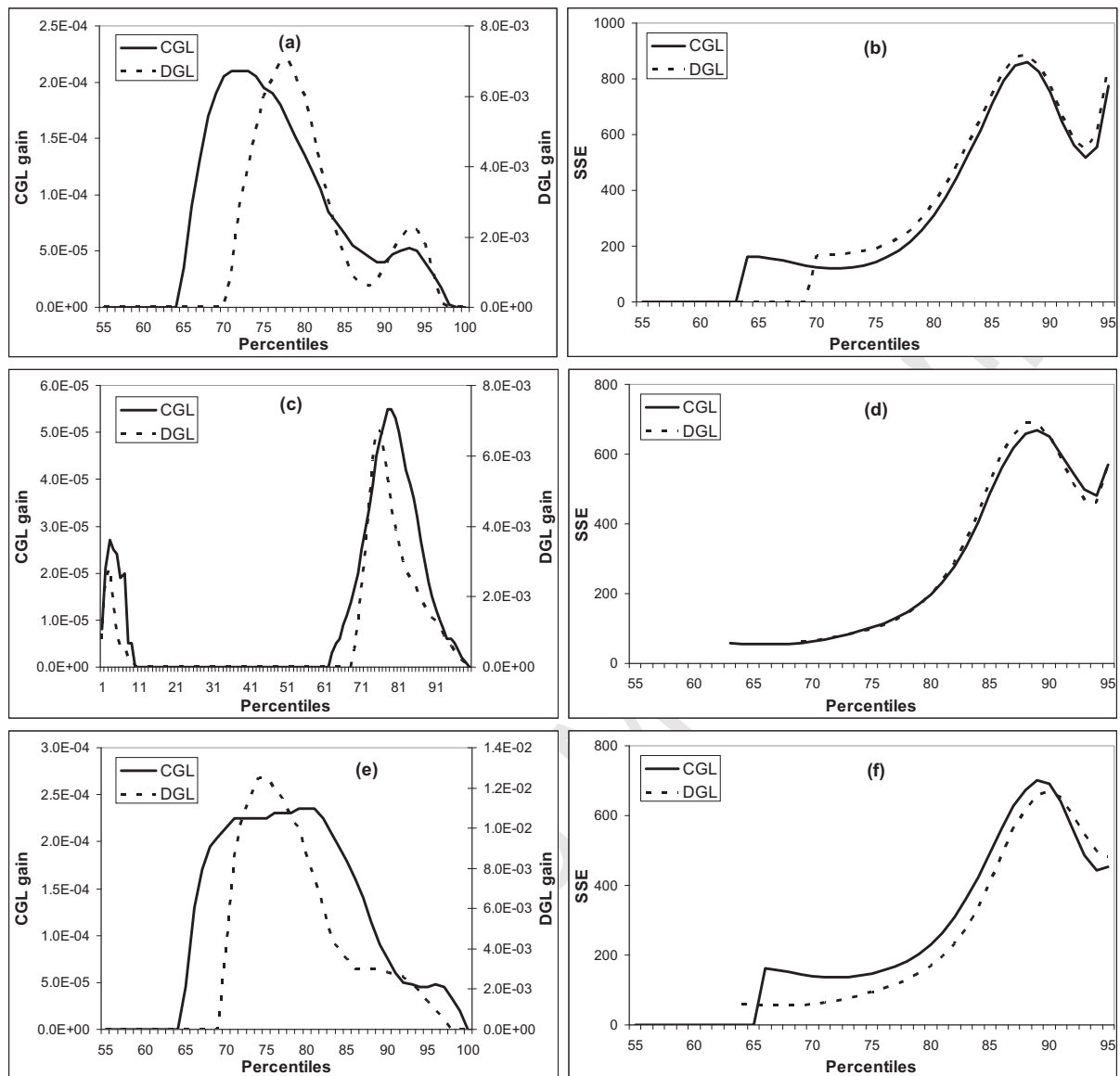


Figure 5: Adaptive learning - standard updating and iterative representation. Notes: for each percentile, the left-hand panel reports the constant and the decreasing gain learning parameter respectively associated with gradient learning and (a) the PLM in Equation (13), (c) the PLM in Equation (14), (e) the PLM in Equation (15). The right-hand panel reports the SSE for each PLM.

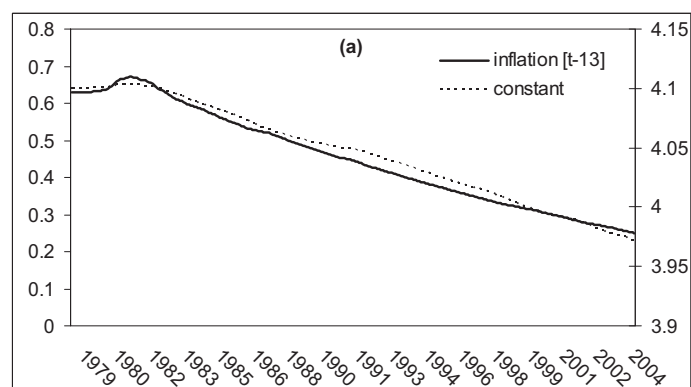


Figure 6: Time-varying coefficients in the PLM  $\pi_{t|t-12} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-13} + \varepsilon_t$  for the 72<sup>nd</sup> percentile - constant gain learning with the standard updating based on past observed errors. Notes: the figure reports  $\hat{\phi}_{1,t-1}$  on the left axis and  $\hat{\phi}_{0,t-1}$  on the right axis.

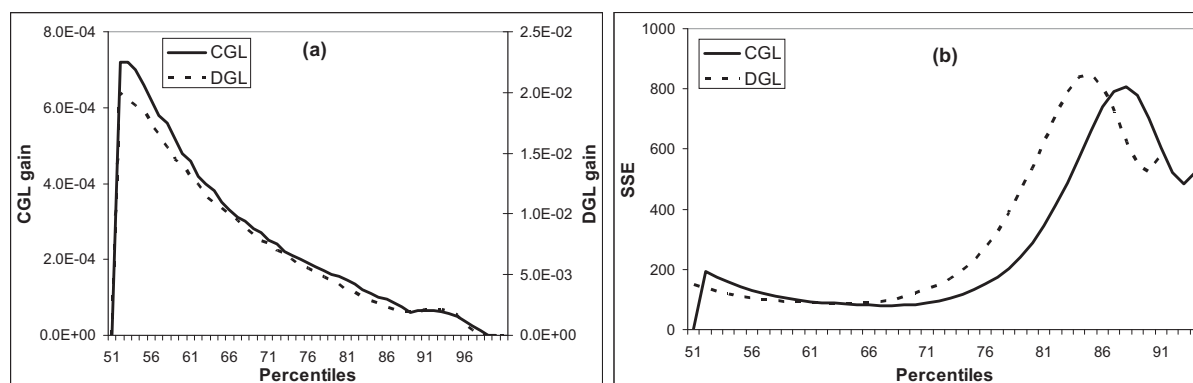


Figure 7: Adaptive learning - updating mechanism based on expected future errors. Notes: for each percentile, Figure 7(a) reports the constant and the decreasing gain learning parameter associated with the updating algorithm in Equation (19). Figure 7(b) reports the associated SSEs.

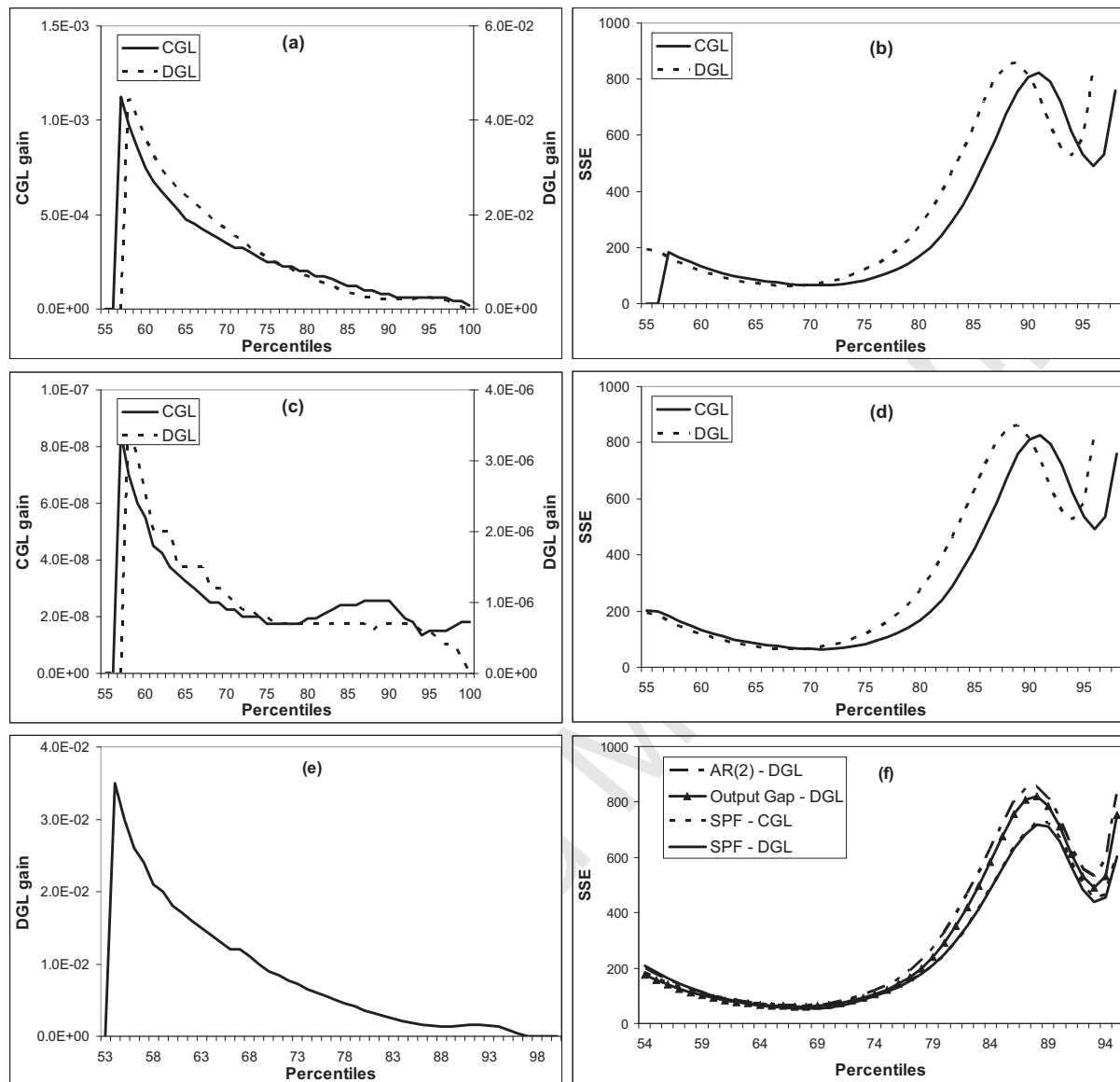


Figure 8: Adaptive learning - updating mechanism based on future errors. Notes: for each percentile, the left-hand panel reports the constant and the decreasing gain learning parameter respectively associated with (a) the PLM in Equation (13) and stochastic gradient learning, (c) the PLM in Equation (13) and least squares learning, (e) the PLM  $\pi_{t|t-12}^s = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1|t-13}^F$  and stochastic gradient learning (only the DG parameter is reported in this case). The right-hand panel reports the SSE for each learning algorithm and the associated PLM. Figure 8(f) reports the SSEs under a stochastic gradient learning scheme and different PLMs that alternatively feature a second lag of inflation (*AR(2) - DGL*), Output Gap (*Output Gap - DGL*), and the SPF forecast under both DGL and CGL (*SPF - DGL* and *SPF - CGL*).

## APPENDIX

Percentile	Constant	$\text{Infl}_{t-13} - \text{Infl}_{t-25}$	$\Delta(\hat{i}_{t-12} - r_{t-12})$	$\Delta y_{t-12}$	$\text{Infl}_t - \text{SPF}_{t-12}$	$\Delta \text{Infl}_{t-12}$	Adj R <sup>2</sup>	DW
5	0.725	0.831	-0.212	-0.012	0.411	0.569	0.913	0.877
	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
	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
	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
	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
	-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
	-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
	-16.674	8.625	1.007	1.583	-3.972	17.933		
	0.000	0.115	0.005	-0.011	-0.087	0.603		

Table A1: Estimation of the model in Equation (4). Notes: a least squares estimation is carried out for each percentile; for each percentile we report the coefficient estimate (first row), the t-stat (second row) and the partial contribution to the adjusted R<sup>2</sup> (third row); all standard errors are corrected for heteroskedasticity and serial correlation using a Newey-West procedure.

Percentile	u	c	$\lambda$	t-test	Percentile	u	c	$\lambda$	t-test	Percentile	u	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: Sticky information - LSTAR model. Notes: we estimate  $\lambda$  in  $\pi_{t|t-12}^k = \lambda F(\pi_{t-13}) \pi_{t|t-12}^F + [1 - \lambda F(\pi_{t-13})] \pi_{t-1|t-13}^k + \varepsilon_t$  (where  $F(\pi_{t-13}) = \{1 + \exp[-v(\pi_{t-13} - c)]\}^{-1}$ ) through a least squares regression while running a grid search on  $v$  and  $c$ , so as to find the combination of parameters that minimize the SSE for each percentile. A t-test is reported for each percentile estimates.

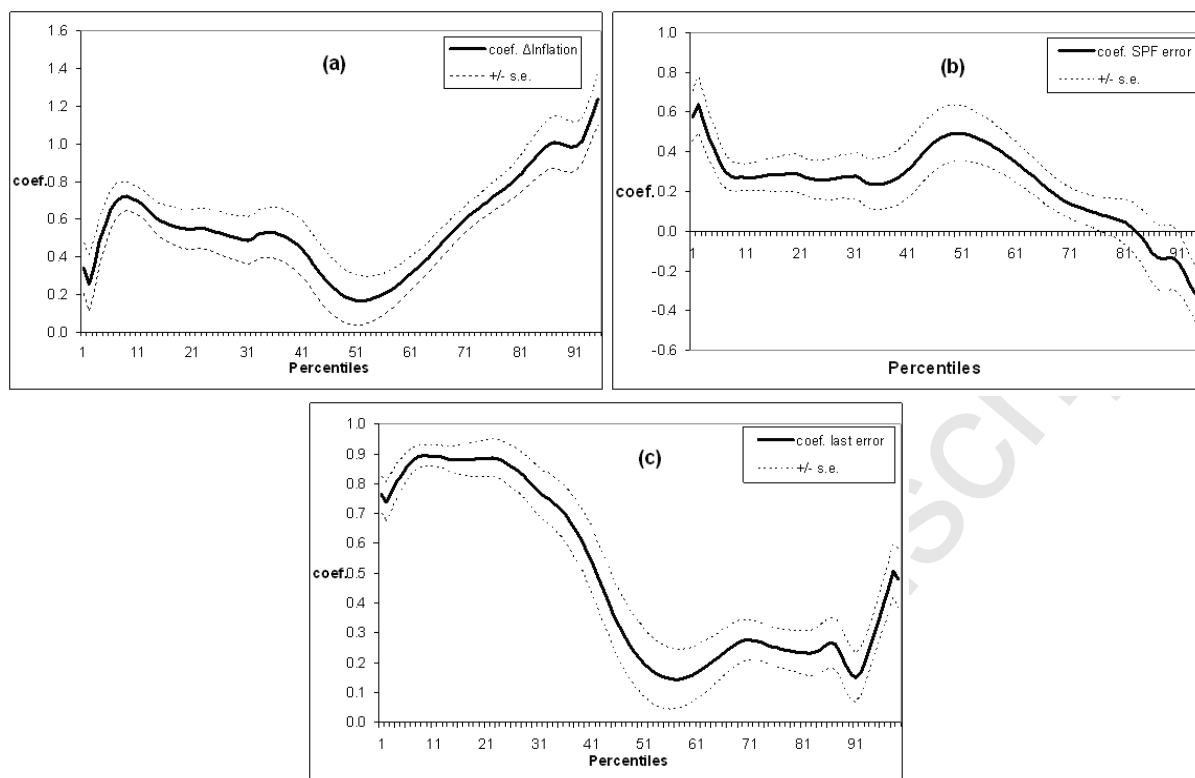


Figure A1: Estimated coefficients from the model in Equation (4). Each graph reports the estimated coefficients associated with: (a)  $\Delta\pi_t$ , (b)  $\pi_t - \pi_{t|t-12}^F$ , (c)  $\pi_{t-13} - \pi_{t-13|t-25}^k$ , respectively. Notes: a least squares regression is run for each percentile; the dotted lines are two standard error bands for the estimated parameter; all standard errors are corrected for heteroskedasticity and serial correlation using a Newey-West procedure.