Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment*

Damjan Pfajfar†
University of Cambridge

Emiliano Santoro‡
University of Cambridge
and University of Trento

3rd August 2006 (Preliminary)

Abstract

Relying on Michigan Survey' monthly micro data on inflation expectations we try to determine the main features – in terms of sources and degree of heterogeneity - of inflation expectation formation over different phases of the business cycle. Different learning rules have been applied to the data, in order to test whether agents are learning and whether their expectations are converging towards perfect foresight. Results suggest that behaviour of agents in the right hand side of the distribution is more associated with learning dynamics. Tests for "static" and "dynamic" versions of sticky information are also conducted. Only agents in the middle of the distribution are regularly updating their information sets. Evidence of rational inattention has been found for agents comprised in the upper end of the distribution. We identify three regions of the overall distribution corresponding to different expectation formation processes, which display a heterogeneous response to main macroeconomic indicators: a static or highly autoregressive (LHS) group, a "nearly" rational group (middle), and a group of agents (RHS) behaving in accordance to adaptive learning and sticky information. The latter, generally speaking, are too "pessimistic" as they overreact to macroeconomic fluctuations.

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 participants at WEHIA 2006, Bologna and Learning Week at St. Louis FED for their comments. All remaining errors are ours.

†Email: dp316@cam.ac.uk. Address: Trinity Hall, Trinity Lane, CB2 1TJ Cambridge, UK

‡Email: es356@cam.ac.uk. Address: Clare Hall, Herschell Road, CB3 9AL Cambridge, UK

Introduction

Throughout the history of economic thought expectations formation process has attracted much attention, although few studies have focused on empirical or experimental evidence on expectations formation process. Several different models have been proposed in the theoretical literature on expectations, but only few of them have been tested, although survey data on household inflation expectations have been now available for decades. Starting point in the theoretical literature is represented by frameworks which assume that expected future values of a variable is equal to the level of the last observation. The first explicit analysis of this expectation rule (usually referred to as naive or static expectations) is due to Ezekiel (1938). The idea of adaptive expectations originates from the work of Fisher (1930) and was formally introduced in the 1950s by several authors, e.g. Nerlove (1958). Nerlove, Grether and Carvalho (1979) first modelled expectations as an autoregressive model of the variables of interest and termed them as quasi-rational expectations. The concept of rational expectations was first discussed in Muth (1961) and in the 1970s it has been popularised by the work of Lucas and Sargent. Lately, a new view of expectations has emerged, postulating that agents act as econometricians when forecasting. This adaptive learning approach is discussed in Evans and Honkapohja (2001).

As far as the empirical literature is concerned the only contributions have come from the introduction of rationality tests and evaluation of adaptive expectation models (Pesaran, 1985, 1987) and, only recently, by an empirical investigation of the degree of heterogeneity (Branch, 2004, 2005) and information stickyness (Mankiw, Reis and Wolfers, 2003 and Carroll, 2003a,b). There have been a few studies that support the introduction of heterogeneous expectations in economic models, e.g. in a standard animal economics model (Baak, 1999 and Chavas, 2000) and in a New Keynesian macroeconomic model (Pfajfar, 2005). Orphanides and Williams (2003, 2005a,b) and Milani (2005a,b) have provided some empirical support for learning dynamics.

Economists know very little about how agents form their expectations in reality. Recently, it can be said that a consensus has been reached on the view that formation is heterogeneous across agents. However, little has been done to investigate formation of expectations in the empirical literature. The studies by Branch (2004, 2005) and Carroll (2003a,b) are noteworthy exceptions. As not all agents have the knowledge of economists we are focusing our research on household survey of inflation expectations. Using monthly micro data on inflation expectations provided by the University of Michigan Survey Research Center we are trying to fill this gap in the literature by attempting to determine the sources of heterogeneity and asymmetries of households' inflation expectations.

There are three main sources of heterogeneity that have been proposed in the literature. Agents might make heterogeneous forecasts because they are relying on different models (different underlying assumptions about structure of the economy, parametrisation, priors), they may have different information sets or they may have different capacities for processing information. Branch (2004, 2005) assesses the importance of the first two sources of heterogeneity and finds that data are consistent with both of them, as the respective models are capable to replicate some characteristics of the empirical distribution. Nevertheless, he concludes that dynamic model uncertainty and dynamic sticky information model deliver better fit than their static counterparts[1]. However he does not consider models which combine both sources of heterogeneity. Carroll (2003a,b) focuses on information constraints as a source of heterogeneity and proposes an epidemiological framework to study how households model inflation expectations. He finds that the diffusion process is rather slow, although the gap between household and professional forecasters narrows down when inflation matters[2] and households become more attentive. In comparison with previous studies, this paper especially focuses on learning and informational stickyness as possible roots of heterogeneity.

We provide evidence that higher moments are important (contrary to Jonung, 1981) for studying expectation formation and also convergence. We find that the cross sectional variance of inflation expectations is counter-cyclical, i.e. it increases during recessions and decreases during booms. However cross sectional skewness and kurtosis are pro-cyclical, both decreasing in recessions and increasing in expansionary periods. Also in the period of stable inflation the variance is less volatile, while skewness and kurtosis are more volatile. We also found some support for convergence lately.

As the pseudo panel we employ is highly unbalanced, we compute percentiles of the empirical distribution, obtaining monthly time series which entail information on the individuals comprised in different parts of the distribution. We perform several tests of rationality, learning, information stickyness and convergence. We find that we cannot reject the hypothesis of rationality just for a few percentiles around or slightly above the median. Test for information stickyness suggest that this is an important source of heterogeneity and that only less than 10% of population are updating their information sets regularly. We also introduce the test for a dynamic version of sticky information model as agents are more likely to regularly update their information sets when inflation "matters". This is found to be a plausible explanation for the center-right hand side of the distribution, as there is significant higher attentiveness in periods of higher inflation. We specially focus on different versions of tests for learning which suggest that agents on the right hand side (RHS) of the distribution tend to behave in an adaptive manner, whereas agents on the left hand side (LHS) of distribution do not exhibit such behaviour. Agents on the RHS of the distribution are particularly associated with updating their coefficients

[1] In this context dynamics is contemplated through the introduction of a Brock and Hommes (1997) switching mechanism between two or more updating frequencies (models).

[2] Carroll proposes that agents update information more frequently when inflation matters due to the increased media coverage on this issue.

with respect to "new information" and slightly less to updating their coefficients with respect to past errors. To further investigate this issue, we estimate several additional time series models of expectation formation. These models confirm a significant degree of heterogeneity and asymmetry in the expectation formation process.

The basic result is that agents positioned around the centre of the distribution behave roughly in line with the rational expectations hypothesis. However, our results suggest that agents on the left-hand side of the distribution behave in an autoregressive way. Furthermore, it can be argued that the inflation expectations of these left of centre agents are stable around some focal points and that they simply do not observe movements in any of the relevant macroeconomic variables. In contrast, on the other side of the distribution, agents are generally too pessimistic and usually produce higher inflation expectations than actual inflation. Curtin (2005) points out that negative changes in inflation have twice the impact as positive changes. As noted above these right of centre agents' inflation expectations are more consistent with adaptive behaviour (learning), although they vary significantly in the speed and method of learning. Furthermore, we argue that they exhibit some features pointed out by recent advances in the macroeconomic and financial literature on inattentiveness and rationally heterogeneous expectations models[3]. We must bear in mind that the cost of being inattentive increases as inflation increases, given that agents have greater incentives to inflation forecasts which entail lower systematic errors.[4] Thus, we carefully study the behaviour of agents over different phases of business cycle.

The remainder of the paper reads as follows: section 1 reports in more detail the dataset employed, while in section 2 we deliver some preliminary descriptive statistics. In section 3 we focus on the percentile time series analysis, with special attention for learning dynamics and informational stickyness. Last section concludes and gives some suggestions for further research.

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, is available at a monthly frequency

[3]Inattentiveness - agents upadate their information sets only occasionally - was advanced by Sims (2003, 2006) and first implemented in macroeconomic model by Mankiw and Reis (2002). The theory of Rationaly Heterogeneous Expectations was put forward by Brock and Hommes (1997). Their basic argument is that it might not always be optimal from utility maximisation point of view to use costly-sophisticated predictor that produce lower mean squared error, thus some agents might be better of with slightly worse predictor which is less costly to use.

[4]More specificaly, Bryan and Palmqvist (2005) study the near-rationality of the survey data analysing if households underpredict inflation when inflation is low. They could not confim the presence of such behaviour for Michigan data while the evidence for Sweden was inconclusive.

from 1978.01.[5] The survey regards an average of 591 households. Each respondent is interviewed once and then reinterviewed after six months. The sampling method is designed in a way that any given month approximately 45% of prior respondents are interviewed, while the remaining 55% is composed by new households. Two relevant questions concerning inflation expectations are whether households expect prices in general to go up, down or to stay the same in the next 12 months, and to quantify the answer. If the answer is that prices will not change the interviewer must make sure that the interviewee actually does not have in mind a rise in the price level at the same rate as the one perceived at the time of the interview.

Although we are aware of the existence of precise quantitative data regarding each respondent and her demographic characteristics, the publicly available version of the survey reports data summarised in groups ("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". Here we follow Curtin (1996) suggestion to regard this answer as 0%. When households expect prices to go up we redistribute the respondents across the six discrete ranges (which predict the price increase), depending on their respective relative shares. We exclude "don't know" respondents from our sample.

As agents are reporting inflation forecasts without any bounds we have to determine the points at both ends of the distribution beyond which observations should be truncated.[6] Curtin (1996) suggests two alternatives: truncation at -10% and +50%, and truncation at -5% and +30%. The two alternatives yield nearly identical trend information, as they are correlated at 0.999%. Overall, there seems to be poor evidence supporting the choice of a truncation rule over the other. The resulting means differ only marginally, and neither truncation rule yields desirable estimates of dispersion. Thus in the following analysis we rely on the smaller truncation range.

2 A preliminary look at the data

In this section we preliminary analyse the available data of the Consumer Survey on Inflation Expectations (CSIE hereafter) between 1978.01 and 2005.02. The second subsection will be devoted to the analysis of the cyclical pattern of the moments of the distribution of inflation expectations. In order to take into consideration different inflation regimes, we will pursue a parallel investigation by considering two subsamples, pre and post 1988.12. This choice allows us to take in adequate account the high inflationary period characterising the first part of the sample.

[5] For more detailed description of the Survey and truncation methods employed see Pfajfar and Santoro (2006).

[6] It is important to note that only estimates of mean and variance of the response distribution are influenced by the exact specification of the truncation rule, whereas estimates of the median of the distribution are unaffected. Technical considerations regarding the cut-off procedure are outlined in Curtin (1996).

2.1 Descriptive statistics

We now perform a brief graphical analysis of the variables of interest. It is worth pointing out at this stage that all the series of expectational variables are plotted at the realised date and not at the date when they were made. Figure 1 plots mean and median against actual inflation.

Insert Figure 1 about here

It is evident how both constantly underestimate the rise in inflation in the first part of the sample, although the forecasting performance improves remarkably during the subsequent deflation. This improvement is probably due to the credibility that the FED acquired in lowering inflationary pressures and, as pointed out before, to a higher opportunity cost of being inattentive in this period. As regards the post-1988 subsample, expectations are quite stable, although they almost systematically fail to forecast periods of low 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, but with opposite sign. On average, it would seem that the mean is a better predictor than the median, although this contradicts the evidence described in Table 1, where the median is always closer to the average inflation. The role of outliers is crucial in the explanation of this situation. In order to shed light on this puzzle we argue that it is useful to conduct a comparative analysis, by splitting the sample in pre and post 1988. This threshold constitutes the end of the cycle and of a marked process of disinflation (see Figure 9), characterised by highly volatile inflation. Figure 6 reports the plot of skewness and kurtosis against the cycle and actual inflation. It is possible to observe that their value is fairly stable and low during the high inflationary period, while it increases and becomes more volatile in the second part of the sample, when inflation is under control. Higher kurtosis and higher positive skewness suggests a higher number of outliers in the right tail of the overall distribution. It would seem at odds that an higher number of outliers arises when inflation is under control. However, following the argument underlying the mechanism developed by Brock and Hommes (1997), it is likely that the opportunity cost of being inattentive or relying on a simple forecasting rule (characterised by a lower degree accuracy) is higher when inflation is high and highly volatile than in periods in which inflation is kept under control. In addition we could argue that in periods of stable inflation there is less media coverage reporting this problem, which translates into a rise in the cost of updating the information set. Furthermore, the prediction accuracy, both in terms of mean and median, is lower during the second part of the sample.

Insert Figure 2, 3, 4 about here

Figure 2 and Figure 3 plot mean, variance, skewness and kurtosis, respectively. It is trivial to observe how higher inflationary expectations are associated to higher volatility. As we already noted in the previous subsection opposite evidence holds with respect to skewness and kurtosis. The data confirm a lower level of skewness and kurtosis in the first part of the sample (opposite evidence holds for the second moment).

Figure 4 reports the 25^{th}, the 50^{th} and the 75^{th} percentiles. This graph helps to understand the different variability characterising different parts of the distribution. In the next section we analyse the macroeconomic determinants of the dynamics of each percentile, in order to detect sources of asymmetry in the response of the distribution over the business cycle. At this stage we limit ourselves to a mere graphical analysis. The 75^{th} percentile appears to be remarkably stable after 1988 with respect to the two remaining series, although the 50^{th} percentile appears to react less and with a marked delay, to the inflationary pressures brought by the first Gulf War, probably because respondents comprised in this range have partially internalised that the rise in inflation is not entirely due to the Gulf War. The 25^{th} percentile, on the contrary, reacts less to the 9/11. Interestingly, the 50^{th} percentile is the more reactive to the 9/11. Perhaps they perceived the thread of deflation as credible at the time.

Insert Figure 5 about here

Figure 5 reports a plot of the mean of the distribution against the actual level of inflation and the mean of the Survey of Professional Forecasters (SPF hereafter) on inflation expectations: it is striking how the former, generally more accurate in the second part of the sample, is more biased than the consumers' one during the highly inflationary period. The two predictions are remarkably similar from 1984 to 1990 and from this point onwards the SPF seems to provide a more accurate prediction.

2.2 Cyclical behavior of the CSIE distribution

In this section we will outline the cyclical features of the empirical moments of the CSIE distribution.

Insert Figure 6, 7 about here

Figures 6 and 7 report the higher moments of the distribution against the output gap series and an indicator of the cycle (the HP detrended Industrial Production Index (IPI) and interpolated estimates of Kuttner' (1994) model of multivariate Kalman filtering). It is clear how variance has a counter-cyclical behaviour, while skewness and kurtosis are highly pro-cyclical. As pointed out before, the third and the fourth moments display higher variability in the post-1988 period. Furthermore, kurtosis exhibits increasing variability in correspondence with the occurrence of peaks in the cycle, which probably reflect uncertainty about the future and hence more unstable tails of the overall distribution.

The dynamics of skewness is strikingly similar to the one chracterising kurtosis: their level points out the existence of a long right tail chracterised by high variability. However, bare eye can lead to conclusion that high peaks in variability are not associated to any cyclical phase or any change in the cycle.[7]

2.3 Time series analysis of the empirical moments

To further investigate the properties of empirical moments we estimate the following two models:

$$em_{t|t-12} = \alpha + \beta t + \gamma \mathbf{X}_t + u_t^{em}, \tag{1}$$

$$\Delta em_{t|t-12} = \gamma \Delta \mathbf{X}_t + u_t^{\Delta em}, \tag{2}$$

$$\mathbf{X}_t^{'} = \begin{bmatrix} y_{t-12} & \pi_{t-12} & i_{t-12} & r_{t-12} & em_{t-1|t-13} & (\pi_{t-12})^2 \end{bmatrix},$$

where y_t denotes a cycle indicator (detrended industrial production index (IPI)), π_t is actual inflation, i_t is the real short term interest rate (3 months t-bill), r_t is the long term interest rate (10 years t-bond yield). We estimate the above equations (em) for the inerquantile range, variance, skewness and kurtosis.

Insert Table C1,C2 about here

We will briefly discuss the results of the analysis on the macroeconomic drivers of the empirical moments of the overall distribution of inflation expectations. The models employed have been reported by equations (1-2). The set of regressors comprises a constant, a deterministic trend, a cycle indicator, actual inflation, squared actual inflation and real short and long term interest rates.

We can confirm the evidence of a negative trend in the variance and of a positive one in skewness and kurtosis, which was clear from a graphical inspection. The interpretation of the estimated coefficients for cycle and inflation at the time of the forecast do not pose any particular problem. At the light of recent theoretical and applied literature on "disagreement" on expected inflation, special attention must be drawn on the evidence we provide on the degree of dispersion. In line with Mankiw, Reis and Wolfers (2003), we want to understand whether different inflation expectations can actually reflect disagreement in the population, and not just mere uncertainty. That is, different forecasts reflect different expectations. Llambros and Zarnowitz (1987) argue that disagreement and uncertainty are two different concepts: intrapersonal variation in expected inflation reflects uncertainty, while interpersonal variation can be conceived as disagreement. They find while there are pronounced changes through time in disagreement, uncertainty varies very little. In order to estimate the degree

[7]This is studied more in depth later when we perform a percentile time series analysis.

of disagreement, Mankiw, Reis and Wolfers (2003) perform an empirical analysis based on Michigan data, Livigstone data and SPF series, finding that the disagreement about the future path of inflation tends to rise with inflation, especially when it changes sharply and in either direction. However, Curtin (2005) argues that increases in the variance occur instantaniously while decreases take place over a longer period of time. Furthermore, it rises in concert with dispersion in rates of inflation across commodity groups and show no clear relationship with measures of real activity. Our evidence points out a marked counter-cyclical behaviour (at the time when forecasts were made), while the coefficient on inflation confirms the results carried out by Mankiw, Reis and Wolfers (2003).

As regards skewness and kurtosis, in accordance with what graphically observed, we can confirm a clear pro-cyclical pattern. On the other side, as we would expect on theoretical grounds, the sign of the estimated coefficients is inverted in the case of inflation. This evidence is particularly important for the rational inattention perspective. Higher and more volatile inflationary pressures should lead agents to raise the level of attention and accuracy in forecasting, while periods of relatively stable inflation, such as the post-1988 period, imply a lower level of attention. A decreasing number of outliers (i.e. lower kurtosis) as inflation increases might fit within this framework. The evidence on kurtosis, moreover, will be confirmed by the percentile regressions models that will be presented in the next section, as units in the upper end of the distribution seem to be more reactive with respect to inflation dynamics. In order to compare the different impacts of the exogenous regressors under different inflationary regimes we have split the sample in pre- and post- December 1988.[8] The empirical exercise shows that coefficients attached to cycle and inflation keep the same sign, although they both decrease in absolute value in the second part of the sample. The interpretation of this evidence can be enriched by adding the effect brought by interest rate regressors. After 1988 agents probably understand the informational content of interest rates: the short term interest rate being the main intermediate target adopted by the FED to fight inflation, while the t-bond yield incorporates a premium for inflation, hence providing an important benchmark for inflation forecasting.

The regressions carried out on the full sample deliver interesting results as regards the informational content of short and long term interest rates. The sign of the estimated coefficients is consistent with the considerations outlined in the previous paragraph. For the variance, t-bill assumes a positive coefficient while t-bond has a negative estimated impact, meaning that if the long term yield increases, it is likely to reflect increased inflation expectations that cause more volatility at a cross sectional dimension, while if the t-bill rate increases, it reflects the will of the central bank to fight inflation strenuously. As regards skewness and kurtosis, the impact of the estimated coefficients is now inverted: higher t-bill rate (which is likely to reflect commitment to fight inflation) leads to an increased number of outliers, especially on

[8] Results are reported in Appendix C.

the RHS of the distribution, probably because agents, relying on the central bank's commitment, have a lower degree of attention. Opposite arguments hold for the estimated coefficient of t-bond yield. In order to check whether a term structure effect is at work, we have also estimated the two models adopting as a regressor the spread between long and short term interest rates. As expected, in the first model the coefficient is highly significant and negative for the third and the fourth moment, while it is positive for the variance. However, the contribution of the change in the horizontal spread to changes in the moments is null. The evidence on the role played by interest rates and the term structure gains more relevance if we consider the post-1988 period, when the new route undertaken FED has been designed by giving more weight to detrimental effects that inflation has. In the first part of the sample the estimated coefficients attached to the interest rate variables are always insignificant in the model expressed in first differences for all the moments under scrutiny. On the other hand for the first model they are significant, although lower in absolute value with respect to their counterparts estimated in the first part of the sample.

3 Percentile Time Series Analysis

In this section we perform a quantile time series analysis. The aim is to move a first step towards the detection of heterogeneity in the response of different regions of the CSIE distribution with respect to macroeconomic variables which are relevant to the rational process of expectation formation. The most important variables are output gap, actual inflation, short and long term interest rates. Furthermore, we introduce in the set of regressors the mean of the distribution determined by the SPF, currently conducted by the FED of Philadelphia[9]. This choice is motivated by the need to observe whether a diffusion process is at work: such a mechanism is likely to have an asymmetric effect on different households. Carroll (2003a,b) designs an epidemiological framework to model how respondents to the Michigan Survey actually form their expectations. For this purpose, he models the evolution of inflationary expectations relying on the assumption that households update their information set from news reports, which at the same time are strongly influenced by the expectations of professional forecasters. As Pesaran and Weale (2005) point out, the diffusion process is, however, slow due to inattentiveness of the households[10].

The choice of the percentile time series, apart from being a useful device to capture

[9] From 1968 to 1990 NBER and ASA were responsible for the conduction of the survey. Before 1981 data exist only for GDP deflator forecasts: we rely on these data for the first few years of our sample.

[10] Thus as Pesaran and Weale (2005) point out, even if the expectations of professional forecasters are rational the expectations of households will only slowly adapt. Carroll (2003a,b) finds that the Michigan Survey has a mean square error on average almost twice that of the SPF. He finds that the SPF inflation expectations Granger-cause household inflation expectations but that opposite evidence does not hold.

asymmetric responses in the distribution of inflationary expectations, is also driven by a more practical consideration: the panel retrievable from the survey is highly unbalanced, as the households interviewed change over time. Thus we extract a set of time series that can be used to capture the evolution of the cross-sectional distribution over the cycle. Here is a brief explanation of the technique employed. Regard the expected change in price level during the following 12 months as a random variable, denoted by π^e, which is distributed with respect to some continuous distribution, $F(\cdot)$. The p^{th} quantile of the distribution, denoted by π^e_p, is the value below which $(100p)\%$ of the distribution lies, hence $F(\pi^e_p) = p$. Thus, we can compute a set of ordered statistics for each month, obtaining $99 (= p)$ time series of percentiles. Of course, the number of observation in the cross-section varies over time. This method is a convenient way to build up a balanced panel of quantiles, after fixing p.

Given our sample sizes, at each cut-off, we can be confident that the estimated quantiles are good estimates of population quantiles: for any two sample ordered statistics $\pi^e_{t,k}$ and $\pi^e_{t,k+h}$, the amount of probability in the population distribution contained in the interval $(\pi^e_{t,k}\ ,\ \pi^e_{t,k+h})$ is a random variable, which does not depend on $F(\cdot)$. Relying on these considerations, for each time period the cross sectional sample is classified in percentiles, thus obtaining 99 time series of percentiles. Percentiles have been obtained by interpolating the distribution obtained after applying the redistribution and the truncation methods outlined in the previous section. Interpolation is a convenient way for obtaining the percentiles at this stage, as the survey reports the percentage of respondents in each range of price movement, hence constituting already a sort of ordered statistic. Furthermore, in order to perform some robustness analysis, different interpolation methods have been applied (such as linear and cubic), which do not yield to major differences.

It needs to be pointed out that strictly speaking we do not have a panel (we have a pseudo panel) since all agents are interviewed only twice. Thus we have a series of cross sections which we treat as a panel. But we can argue that there is some support that agents with similar characteristics behave similarly. As we know that agents have to forecast inflation on a daily basis when they make economic decisions and not only when we ask them to do so. Thus we could put forward some arguments in line with overlapping generations models. Furthermore, we have to point out that strictly speaking we consider a "representative agent" for each percentile, which cannot move across percentiles, although we found bellow some evidence that this restriction is likely to be violated in the data[11]. We acknowledge that these restrictions are rather strong, but we still feel that we can retrieve some valuable information about inflation

[11] Also Branch (2004, 2005) has provided some support for time-varing degrees of heterogeneity. However, Curtin (2005) reports that 73% of agents in their second interwiew for CSIE update their inflation expectations while 27% report the same inflation expectations as in the first interview (for 1993-2005 period). 35% of agents reported higher inflation expectations in the second interview and 38% reported lower inflation expectations. The monthly change in actual inflation in this period was 0.0005 percentage points.

expectations process despite these unavoidable restrictions that are present in the data.

3.1 Rationality tests

We now apply some tests of rationality commonly employed in the literature[12]. The rational expectations hypothesis can be interestingly applied to survey expectations data, as these allow to determine different degrees of forecast efficiency. The latter has to be intended as the result of a forecasting procedure that does not yield to predictable errors. The simplest test of efficiency is a test of bias[13]. It is possible, by regressing the expectation error on a constant, to verify whether inflation expectations are centred around the right value:

$$\pi_t - \pi^k_{t|t-12} = \alpha + \varepsilon_t, \tag{3}$$

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

$$\pi_t = a + b\pi^k_{t|t-12} + \varepsilon_t, \tag{4}$$

where rationality implies that $a = 0$ and $b = 1$, jointly. The last expression can be simply augmented to test whether information in a forecast is fully exploited:

$$\pi_t - \pi^k_{t|t-12} = a + (b-1)\,\pi^k_{t|t-12} + \varepsilon_t. \tag{5}$$

Testing remains the same as in the previous regression: under the null of rationality these regressions are meant to have no predictive power[14].

Results When running regressions on equation (3) we can observe that only agents between the 51^{st}-55^{th} (52^{nd}-54^{th}) percentile range are not biased at 1% significance (5% significance) level. Test of biasness have been conducted many times on different survey data. Croushore (1998), Roberts (1997) and Mankiew, Reis and Wolfers (2003) have among others studied rationality of the mean and/or the median of Michigan data and found that they can almost always reject the null of rationality. As we have seen there are some percentiles slightly above the median for which the null of rationality could not be rejected. When splitting the sample into pre-1988 and

[12] See Pesaran (1989), Mankiw, Reis and Wolfers (2003) and Bakhshi and Yates (1998) for a review of these tests.

[13] See, for an application, Jonung and Laidler (1988) and Mankiw, Reis, and Wolfers (2003).

[14] Alternative test for rationality could take into account that inflation and inflation expectations data are I(1): in this situation the rational expectations hypothesis suggests that they cointegrate, i.e. that expectations errors are stationary, and that the cointegrating vector has no constant terms, as well as coefficients on expected and actual inflation, which are equal in absolute value (Bakhshi and Yates, 1998).

post-1988, we find that for the pre-1988 sample agents between the 55^{th}-63^{rd} (56^{th}-62^{nd}) percentile are not biased at 1% significance (5% significance) level. For the 1999-2005 period we found that agents between the 47^{th}-50^{th} (48^{th}-50^{th}) percentile are not biased at 1% significance (5% significance) level. We can observe that there are more rational agents in the first subsample when inflation was higher and was probably more important to produce better forecasts. By estimating equation (5) and computing the F-test we find that it is always possible to reject the null hypothesis (rationality) that the first coefficient (a) is 0 and the second (b) is 1 for the whole sample and the two sub-samples. Similar conclusion have been reached by the studies mention above when analysing the mean and the median of the CSIE.

3.2 Learning

3.2.1 Estimating Simple Learning Rules

In the next sections we investigate the importance of adaptive behaviour of agents. Different learning rules will be implemented for the Michigan Survey data, in order to test whether agents' expectations are converging towards rational expectations (perfect foresight). For a discussion on different learning rules and convergence to rational expectations see Evans and Honkapohja (2001). Learning will be first tested in a model with constant gain learning, where convergence to rational expectations is not generally observed. The model below is equivalent to the adaptive expectations formula:

$$\pi^k_{t|t-12} = \pi^k_{t-13|t-25} + \vartheta \left(\pi_{t-13} - \pi^k_{t-13|t-25}\right) + \varepsilon_t, \tag{6}$$

where ϑ is the constant gain parameter. Under this learning rule agents revise their expectations according to the error of the last realised forecast. Since in the survey of inflation expectations agents are asked to forecast inflation in the next year time (hence they make their forecast at time $t-12$), the revision will regard the previous period's forecast (at time $t-13$), which was made at time $t-25$.

Below we represent a learning mechanism with decreasing gain parameter:

$$\pi^k_{t|t-12} = \pi^k_{t-13|t-25} + \frac{\iota}{t^{\varkappa}} \left(\pi_{t-13} - \pi^k_{t-13|t-25}\right) + \varepsilon_t. \tag{7}$$

The empirical approach will consist in estimating ϑ and ι. $\varkappa$ is the coefficient that controls the dampening of the learning gain. If we want that the learning always converges to the rational expectations $\varkappa \leq 1$. If the estimated parameters will be significantly different from 0, then we could conclude that agents are actually learning from their past mistakes.

Results We start by analysing the degree of adaptiveness of inflation expectations by estimating equations (6) and (7).[15]

[15] Results are oulined below in the Table 8.

Insert Figures 8-11 about here

Our estimates suggest that agents in the upper part of the distribution are behaving at least partly in an adaptive manner while for the agents comprised in the poor hand of the distribution the past error has little or no explanatory power. With regards to the estimated constant gain and the overall R^2, we can observe a hump-shaped response between the 40^{th} and the 99^{th} percentile which peaks around the 75^{th} percentile, i.e. in the range in which percentile past errors have the highest explanatory power. Below we will generalise this regression by including other possible explanatory variables.

The decreasing gain learning estimates confirm that agents between the 40^{th} and 95^{th} percentile are behaving partly in an adaptive way. Indeed, the decreasing gain estimates suggest that this method of learning is more in line with the bahaviour of agents in the upper part of distribution. As noticed before, this method of learning has no explanatory power for agents in the left-hand side of the distribution. Also in this case we observe a hump-shaped response, although the adjusted R^2 peaks around 0.75, compared to a value of about 0.35 obtained in the case of constant gain learning. A higher explanatory power of the decreasing gain learning might be due to the high inflationary period at the beginning of our sample.

3.2.2 Recursive Representation of Simple Learning Rules: first version

The above specification is mainly aimed at testing whether data support the existence of adaptive behaviour. As in the adaptive learning literature it is assumed that agents behave like econometricians using all the available information at the time of the forecast, we have to specify a recursive model for the two different learning rules mentioned above. In this version we are testing if agents are updating their coefficients with respect to the last observed error. We will assume that agents' perceived law of motion (PLM) will be a simple AR(1) process[16]

$$\pi^s_{t|t-12} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-13} + \varepsilon_t. \tag{8}$$

When agents are estimating their PLM they exploit all the available information up to period $t-1$. As new data become available they update their estimates according to a constant gain learning (CGL) or a decreasing gain learning (DGL) rule. First we specify stochastic gradient learning with constant or decreasing gain and then we focus on least square learning. Let X_t and $\widehat{\phi}_t$ be the following vectors: $X_t = \begin{pmatrix} 1 & \pi_t \end{pmatrix}$ and $\widehat{\phi}_t = \begin{pmatrix} \phi_{0,t} & \phi_{1,t} \end{pmatrix}'$. In stochastic gradient learning (see Evans, Honkapohja and Williams, 2005) agents update coefficients according to the following rule:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-25}\left(\pi_{t-12} - X_{t-25}\widehat{\phi}_{t-13}\right). \tag{9}$$

[16] In the remainder we will analyse several different PLMs.

In the updating algorithm for decreasing gain learning we just replace ϑ with $\frac{\iota}{t}$. In least square learning agents take into account also the matrix of second moments of X_t, R_t. In the case of constant gain they update their coefficients in the following way:

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

$$R_t = R_{t-1} + \vartheta \left(X_{t-25} X'_{t-25} - R_{t-1} \right). \tag{11}$$

As before, in the updating algorithm for decreasing gain learning we replace ϑ with $\frac{\iota}{t}$. The empirical approach will consist in finding ϑ and ι that minimise the sum of squared errors (SSE), i.e. $\left(\pi^s_{t|t-12} - \pi^k_{t|t-12} \right)^2$. The drawback of this approach is that we have to assume the initial values for $\widehat{\phi}_t$ for 12 periods.

When we are recursively estimating learning the main problem is how to set initial values. This problem is extensively discussed in Carceles-Poveda and Giannitsarou (2005). Strictly speaking this problem should not occur in our case since we are simply trying to replicate our time-series data as closely as possible. Thus in the following recursive learning estimations we design this exercise in order to search for the best combinations of gain and initial values to match each percentile as closely as possible[17]. This version of initialisation could also be considered for testing the series if it exhibits learning, i.e. the gain is positive under this method of initialisation it would exhibits learning for all other initialisation methods with higher (or equal) gain.

Results Agents between the 65^{th} and the 98^{th} percentile are behaving in accordance with constant gain version of gradient learning. The estimated gain can be observed in Figure 12. It can be seen that the gain has a hump-shaped pattern, reaching a peak at $2.1 * 10^{-4}$, which is located between 71^{st} and 73^{rd} percentile. There is also another smaller peak around the 93^{rd} percentile. Decreasing gain version of gradient learning is significant for agents between 70^{th} and 96^{th} percentile. The estimated gain has similar properties to the case of constant gain, with the exception that the second "hump" is slightly more evident in the decreasing gain case (see Figure 13). The highest gain is estimated at $0.007t^{-1}$ for 76^{th} and 77^{th} percentile. To compare both versions of this gradient learning we have plotted SSE for both cases in Figure 14. The results are suggesting that the constant gain version of gradient learning better describes the behaviour of agents, especially around the 70^{th} percentile.

Insert Figures 12-14

Comparing to previous estimates of gain coefficient, e.g. Orphanides and Williams (2005a) suggested 0.01-0.04, Milani (2005a) estimates it 0.0183, our estimates are

[17] This approach has however an obvious practical inconvenient as running a grid search on several variables it is computationally very intensive.

rather small. But their estimates are with quarterly data while ours are with monthly data. To make the results more comparable we have to compute how many past information agents use to form their expectations. An estimate of 0.02 with quarterly data suggests that agents are using 12.5 years of data, while an estimates of $2.1 * 10^{-4}$ with monthly data implies that agents are using roughly 400 years of data. However we have to treat these estimates as lower bound estimates of the gain coefficient as we are searching for optimal values of initial values as well and thus any other method for initialising learning would result in higher gain coefficient.

The results when taking into account the matrix of second moments[18] are very similar to the results just reported as the covariance terms are quite small. Thus we are rather focusing on the results where we augment the agents' respective PLM to the case when they do not include inflation in the previous period in the PLM, but just their forecasts in the previous period. This version of learning is found to better replicate the behaviour of agents than the previous version of PLM. For this version of learning also some agents on the LHS of the distribution are learning. In a CGL version agents between the 1^{st} and the 9^{th} and the 63^{rd} and the 99^{th} percentile are behaving adaptively while in DGL version between the 1^{st} and the 9^{th} and the 69^{th} and the 99^{th} percentile. In the constant gain case the response pattern on the RHS is quite similar to normal distribution, with the highest gain for 78^{th} and 79^{th} percentile ($5.5 * 10^{-5}$). The response pattern for decreasing gain is also hump shaped on the RHS, but it reaches the highest gain for 75^{th} and 76^{th} percentile ($0.0067t^{-1}$). For most percentile constant gain learning better describes the behaviour of agents, except for agents around 75^{th} percentile.

Insert Figures 15-16

3.2.3 Recursive Representation of Simple Learning Rules: second version

In this section we assume a slightly different way of forming expectations. The process for the PLM is again assumed to be an AR(1):

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

We implement the following gradient learning updating algorithm:

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

In the updating algorithm for decreasing gain learning we replace ϑ with $\frac{\iota}{t}$. As we are studying forecasts 12 months ahead, agents will derive 12-months ahead forecasts in the following way:

$$\pi^s_{t+12|t} = \phi_{0,t-1}\left[1 + \phi_{1,t-1} + \left(\phi_{1,t-1}\right)^2 + ... + \left(\phi_{1,t-1}\right)^{12}\right] + \left(\phi_{1,t-1}\right)^{13}\pi_{t-1}. \tag{14}$$

[18] We set the matrix of second moments to be constant and equal to the average in the sample employed in the analysis.

With this approach we have to assume only the initial values for $\widehat{\phi}_t$ for 1 period. As before, the empirical approach will consist in finding ϑ and ι that minimise the SSE.

Results Results reached through this different specification are quite similar to the results in the previous one, although due to the particular way of forming beliefs in this section the results explain slightly less accurately the behaviour of agents. Nevertheless, our estimates suggest that agents between 65^{th} and 99^{th} percentile are behaving in accordance with constant gain version of the gradient learning algorithm. In this case it is even more evident the hump-shaped pattern of optimal gains. The gain peaks at $2.35 * 10^{-4}$ between the 79^{th} and the 82^{nd} percentile (see Figure 19). Also in this version of learning there are less agents associated with decreasing gain learning. For agents between 70^{th} and 97^{th} percentile we have found positive optimal gains with the highest value of $0.0125t^{-1}$ for 74^{th} and 75^{th} percentile (see Figure 20). Another similarity with the previous version is constituted by the fact that the constant gain learning is constantly outperforming the decreasing gain learning (see Figure 21). The SSE reaches a minimum around the 70^{th} percentile.

Insert Figures 17-19

3.2.4 Recursive Representation of Simple Learning Rules: third version

The third version of recursive learning assumes that agents are updating their coefficient estimates with new information that is released about future inflation. So they are updating parameters regarding the "future errors". This version of learning was first advanced in Fuchs (1979). Again we are assuming PLM of an AR(1) form, although we will also test other versions of PLM:

$$\pi^s_{t+12|t} = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1} + \varepsilon_t. \tag{15}$$

First we assume gradient learning updating algorithm:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta X'_{t-1}\left(\pi_{t+12} - X_{t-1}\widehat{\phi}_{t-1}\right). \tag{16}$$

In the updating algorithm for decreasing gain learning we replace ϑ with $\frac{\iota}{t}$. As in the first version we will also explore least square learning, where agents take into account also the matrix of second moments of X_t, R_t. In the case of constant gain they update their coefficients in the following way:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \vartheta R^{-1}_{t-1}X'_{t-1}\left(\pi_{t+12} - X_{t-1}\widehat{\phi}_{t-1}\right); \tag{17}$$

$$R_t = R_{t-1} + \vartheta\left(X_{t-1}X'_{t-1} - R_{t-1}\right). \tag{18}$$

As before, in the updating algorithm for decreasing gain learning we replace ϑ with $\frac{\iota}{t}$. The empirical approach will consist in finding ϑ and ι that minimise the SSE.

Results The two previous learning rules are more associated with learning from past errors, while this version is more "forward-looking" and it is estimating whether agents are updating after new information comes available. The general finding is that majority of agents on the RHS is more associated with this version of adaptive learning algorithm. Hence we will focus more on this version of learning and investigate different potential PLMs for agents. We will start with simple AR(1) form PLM as in the above two versions in order that we can directly compare the results.

Results are suggesting that agents start learning above the 55^{th} percentile for constant gain and at 56^{th} percentile for decreasing gain, where the gain immediately jumps to the highest value. Afterwards the gain starts slowly decaying and converging to zero. At about 99^{th} percentile it converges to 0 at constant gain case and at about 98^{th} for decreasing gain case. As we can see in the below figures the highest gain is about $1.125 * 10^{-3}$ and the lowest SSE is at about 68^{th} percentile. Compared to estimates in the first version this gain is already more "realistic" as it suggests that agent are using about 74 years of data, although still quite high. For decreasing gain learning the highest gain is $0.0445t^{-1}$ while the SSE are very similar to those of CGL. Strictly speaking constant gain case does slightly better for most of the percentile but between 63^{rd} and 69^{th} percentile.

Insert Figures 20-22 about here

Next we focus on least squares learning. We set the variance-covariance matrix to be constant across the percentiles (to the average of the sample) as running the grid search also on initial values of variance-covariance matrix would computationally too intensive. The results in this case are extremely similar to the above case where we do not take into account the matrix of second moments. SSE are practically the same in both cases (see Figure 25), just the optimal gain parameters are different and also the pattern of optimal gains is slightly different (see Figures 23 and 24). Maximum optimal gain is $8.5 * 10^{-8}$ in the case of constant gain learning and $3.5 * 10^{-6}t^{-1}$ in the case of decreasing gain learning.

Insert Figures 23-25 about here

Next we try different PLMs. We estimate learning by adding into PLM second lag of inflation, output gap or inflation forecasts by professional forecasters. We estimate both decreasing and constant gain learning. We find that PLM with inflation forecasts of professional forecasters is performing better than the other options, especially the decreasing gain version of this learning. The response pattern for optimal gain is quite similar for all the PLMs (In the case for decreasing gain SPF it is plotted at Figure 26). With these expanded PLMs agents between 54^{th} and 98^{th} percentile are behaving in the adaptive way. Figure 27 plots SSE for different PLMs and Table 2 reports maximum gains for different PLMs. (Table 2)

Insert Figures 26-27 about here

The results confirm our initial conjecture that behaviour of agents in the RHS of distribution is more closely associated with learning dynamics as specified above. The "optimal" gain in CGL was estimated between 0 and 0.051. Overall we can say that decreasing gain learning better replicates the behaviour of majority of agents.

3.2.5 Testing for convergence: Weighted least squares learning and the Kalman filter

In this section the coefficients in the PLM are updated through the following algorithm:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \frac{\alpha_t}{t} R_{t-1}^{-1} X_t' \left(\pi_t - X_t \widehat{\phi}_{t-1} \right), \tag{19}$$

where $R_t = \frac{1}{t}\sum_{\tau=1}^{t} \alpha_\tau X_\tau X_\tau'$ and α_τ is a sequence of positive numbers. This formula is a version of weighted least squares, which also corresponds to recursive least squares for $\alpha_\tau = 1$. This updating procedure can be implemented within a Kalman filter framework. After substituting $P_t = \frac{1}{t} R_t^{-1}$ and $f_t = X_t P_{t-1} X_t' + \frac{1}{\alpha_t}$ in (19) we end up with:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + P_{t-1} X_t' f_t^{-1} \left(\pi_t - X_t \widehat{\phi}_{t-1} \right);$$

$$P_t = P_{t-1} - P_{t-1} X_t' X_\tau P_{t-1} f_t^{-1},$$

which corresponds to the state-space model:

$$\pi_{t|t-12}^{s} = \phi_{0,t} + \phi_{1,t} \pi_{t-13} + e_t;$$

$$\forall i \quad \phi_{i,t} = \phi_{i,t-1} + \eta_{i,t},$$

with hyper-parameters given by

$$Var\left(e_t\right) = \frac{1}{\alpha_t}; \tag{20}$$

$$Var\left(\eta_t\right) = 0. \tag{21}$$

As the least square estimation assumes that the coefficients are stable, while their estimated counterparts are time varying, the learning process is not optimal. The results reported in Marcet and Sargent (1989) on the convergence of the learning process towards rational expectations only hold when the law of motions for the parameters are viewed as invariant. Hence, if $Var\left(\eta_t\right) \neq 0$ then P_t does not converge towards 0, and consequently the learning does not converge to rational expectations. Under a more general state-space setting coefficients would be derived as follows:

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + P_{t-1} X_t' f_t^{-1} \left(\pi_t - X_t \widehat{\phi}_{t-1} \right); \tag{22}$$

$$P_t = P_{t-1} + Q_t - P_{t-1} X_t' X_\tau P_{t-1} f_t^{-1}, \tag{23}$$

where $f_t = X_t P_{t-1} X_t' + H_t$, $Var(e_t) = H_t$ and $Var(\eta_t) = Q_t$. Therefore, the expectations of bounded rational agents are computed as the prediction of π_t^s:

$$\pi^s_{t|t-12} = \phi_{1,t|t-1} + \phi_{2,t|t-1}\pi_{t-13}. \tag{24}$$

Note that Kalman filter delivers the optimal gain that agents apply when updating their parameters. It also allows to test whether the learning is perpetual or whether is converging to rational expectations. Practically, the procedure implies a test of significance of the variance of the state variables.

Results Following Basdevant (2003) we constrained the variance of each state variable to be identical, since the data set is relatively limited. Initially we estimate a "standard" state-space model where Q_t is regarded as a constant. This can be conceived as a test for permanent learning. Broadly speaking, all agents are learning when we allow for optimal gain in each period. The general cross sectional pattern pointed out in the percentile regressions emerges in this case, i.e. hump-shaped in the central region (20^{th}-70^{th} percentile), with a maximum of 0.1 around the 40th when testing for perpetual learning. There is also a narrower but more peaked pattern in the variance, Q_t, between the 70^{th} and 100^{th} percentile: agents comprised in this region are characterised by a slower (perpetual) learning.

To check whether the hyper-parameter is decreasing over time, a second model has been estimated. Following the ideas discussed in Hall, Robertson and Wickens (1997) on convergence we model $Q_t = Q_0(t)$. A value significantly different from 0 and lower than 1 would imply that the process is converging towards recursive least squares and thus it moves towards a rational expectations equilibrium. The cross sectional pattern suggests that criteria for convergence are fulfilled at every percentile, although the convergence process is generally quite slow, especially in the upper end of the distribution.

Hence, there is little difference between the two models, as the convergence implied by the second one is very slow and the first model implies perpetual learning.

Insert Figures 28-29 about here

3.3 Sticky Information

3.3.1 Testing for Sticky Information - Static Case

In this section we estimate a simple regression introduced in Carroll (2003), in order to investigate the relevance of a static sticky information model for our dataset. We estimate the following equation

$$\pi^k_{t|t-12} = \lambda_1 \pi^s_{t|t-12} + (1-\lambda_1)\,\pi^k_{t-1|t-13} + \varepsilon_t. \tag{25}$$

As Carroll (2003) points out, news about inflation spread slowly across agents, reaching only a fraction λ_1 of population in each period. We will estimate the model with the restriction on coefficients although this restriction is not likely to be satisfied across all percentiles[19].

Results Figure 30 plots, for each percentile, the value λ_1^{-1}, which provides us with an estimate of the average updating period. The estimation confirms the existence of static behaviour in the informational structure up to the 40^{th} percentile. From this point up to the 91^{st} percentile we can detect the presence of a U-shaped pattern, with a minimum occurring at the 50^{th} percentile, which translates into an average minimum updating time of 7 months, lying within a range of 5 to 12 months. Carroll (2003a) found similar results as his estimate of λ_1^{-1} for the mean was 11 months[20], while Döpke et. all (2006) estimate for European data was roughly 18 months. Mankiew, Reis and Wolfers (2003) and Branch (2005) fix $\lambda = 0.1$ which for monthly data implies average updating of 10 months. Branch (2005) further investigates sticky information argument by allowing for switching between different updating frequencies. We estimate heterogeneities in updating frequencies with alternative version of dynamic sticky information in the next section.

Insert Figure 30 about here

3.3.2 Testing for Sticky Information - Dynamic Case

Rational inattention has also another simple testable implication: when inflation "matters" agents will update their information sets more frequently, in order to forecast more accurately. We will assume that a higher proportion of agents will pay attention to new information coming available when inflation is higher as their opportunity cost of being inattentive is significantly higher during these periods. To test this hypothesis we relax the assumption of linearity in the equation (25) and assume a particular non-linearity structure in the form of a logistic Smooth-Transition Autoregressive (LSTAR) model.[21]:

$$\pi^k_{t|t-12} = \lambda_1 F\left(\pi_{t-12}\right) \pi^s_{t|t-12} + \left[1 - \lambda_1 F\left(\pi_{t-12}\right)\right] \pi^k_{t-1|t-13} + \varepsilon_t, \tag{26}$$

where function F is the following logistic function:

$$F\left(\pi_{t-12}\right) = \frac{1}{1 + \exp\left[-\upsilon\left(\pi_{t-12} - c\right)\right]}. \tag{27}$$

The estimation procedure will consist 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

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

[20] Mankiew and Reis (2002) have implemented $\lambda = 0.25$ (average upadating 12 months) in their simulations assuming quarterly data.

[21] For details about Smooth-Transition Regression model see e.g. Granger and Teräsvirta (1993).

minimises the sum of squared errors for each percentile. Some support for the version of dynamic sticky information presented here can be put forward by observing the cyclical behaviour of the moments of CSIE distribution.

Results The estimates of non-linear version of the model reported in equation (26) are very similar to the results of the linear counterpart. The average difference of SSE is only about -0.595. Nevertheless we get positive coefficients in the transition function for all percentiles, so that the non-linear version of the model always outperform linear version, although as mention above only marginally. Responses between 59^{th} and 79^{th} percentile are the most clearly associated with rational inattention argument as their attention is higher in periods of high inflation and lower in periods of low inflation.

Insert Figure 31 about here

Figure 31 plots the value $\left[\lambda_1 F\left(\pi_{t-12}\right)\right]^{-1}$, which is a time-varying estimate of average updating period for 52^{nd} and 63^{rd} percentile. As we can notice, especially the average updating period for 63^{rd} percentile behaves in accordance with rational inattention view. At the beginning of the sample agents are updating information regularly as inflation is higher and thus the opportunity cost of not updating the information set is higher. The optimal coefficients in the transition function for this percentile are $\upsilon = 0.21$ and $\iota = 2.58$. The latter coefficient can be interpreted as perceived implicit inflation target of the Federal Reserve System for our sample. The dynamics of the average updating period for the 52^{nd} percentile is quite different as agents comprised in this percentile are only in certain periods slightly less attentive. The optimal coefficients in the transition function for the 52^{nd} percentile are $\upsilon = 3.18$ and $\iota = 7.40$. As the former coefficient is higher than 1, the interpretation of these coefficients is different from the previous case. The ι cannot be interpreted as perceived inflation target. For these agents the difference between linear and non-linear model is really small. Agents between 50^{th} and 58^{th} percentile have similar responses, but for higher percentiles in this range the variability of the estimated average updating frequency is higher. Thus in periods when inflation "does not matter" they become more inattentive. Similar dynamics is also observed for the range above the 80^{th} percentile, although with a much higher average time to update information. Overall there are three different dynamic patterns that we can observe depending on the optimal coefficients in the transition function (see Table 1). We have already characterised the first two cases, while the third one occurs when υ is low (below 1) and ι is as high as in the previous case (above 5). Also in this case the response pattern is quite similar to the case when υ and ι are high, just the variability is higher in the case when υ is low.

This version of dynamic sticky information is alternative to specification in Branch (2005) where he models it as a Brock and Hommes (1997) type of choice between different updating frequencies. He finds that majority of agents update their information sets every 3-6 months, while less agents update their information sets every period. He find that some agents update their information sets every 9 months or even less frequently. Contrary to these results we provide evidence that average updating frequency are higher at least on the LHS of the distribution, while also agents on RHS although they are behaving in accordance with rational inattention argument are updating on average less frequently than once a year.

3.4 "General" Models of Expectation Formation

3.4.1 What Macro Variables Do Agents Consider to Form Their Expectations?

We also estimate some more general models of expectations formation. The first model investigates which variables agents take into account when forecasting inflation. We specify the following percentile regression:

$$\begin{aligned} \pi^k_{t|t-12} &= \alpha + \sum_i \gamma_i \pi_{t-i} + \sum_i \beta_i y_{t-i} + \mu i_{t-24} + \delta r_{t-24} + \zeta \pi^k_{t-1|t-13} + \eta \pi^F_{t|t-12} + \varepsilon_t \quad (28) \\ k &= 1, ..., 99; \qquad i = 12, 14, 24, 30. \end{aligned}$$

We denote with $\pi^k_{t|t-12}$ 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 derived from the SPF.

Results As already mentioned the model (28) aims at characterising the relevance of the determinants of the one-year-ahead inflation expectations. We introduce in the set of exogenous regressors the following contemporaneous (i.e. at the time in which expectations are formed) and lagged variables: rate of inflation, cycle (measured as the Hodrick-Prescott [HP] detrended Index of Industrial Production [IPI]), short term interest rate (3-month treasury bill), long term interest rate (10-year bond yield) and the one-year-ahead expectation taken from the SPF. It turns out that just some of the mentioned regressors can actually account for movements in the dependent variable and contextually have a clear cut interpretation. Thus, in our model we are setting $\gamma_{14} = ... = \gamma_{30} = 0$ and $\beta_{14} = ... = \beta_{30} = \mu = 0$ as they are almost always insignificant. Usually we find that contemporaneous rate of inflation and the autoregressive term and to some extent the SPF forecast have significant predictive power. As far as the remaining regressors are concerned, on empirical grounds we can argue that these variables are generally either not observed or not taken into account for the determination of the expectation at each range of the CSIE distribution. Only the contemporaneous cycle is marginally significant. The resulting response functions have been plotted in figures in Appendix B. Furthermore, figure 8 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))[22]. This statistics provides important information on the different information structure underlying the mechanism of expectation formation for the individuals comprised in different ranges of the distribution.

Insert Figure 32 about here
Insert Table 3 about here

As it is clear from Figure 32, in the upper tail of the distribution the constant (from the 85^{th} percentile onwards) as well as the estimated coefficient associated to the actual inflation (from the 70^{th} percentile onwards) take high values. This element corroborates the evidence arising from the observation of the descriptive statistics, confirming a marked degree of pessimism for the upper tail of the distribution. On the other hand, looking at the response function for the actual inflation in the middle range (in the interval $[25^{th}, 70^{th}]$), we can notice an evident hump-shaped pattern. However within the same interval the autoregressive term implies a U-shaped response. These results are in line with what we would expect on theoretical grounds, as more rational individuals should rely less on past expectations. They should also display a lower degree of stickyness, and rely more on actual inflation, which is likely to have a higher informational content.

An interesting situation can be outlined from the observation of the graph reporting the overall R^2 and the partial "contribution" coefficients. It is clear that up to the 70^{th} percentile most of the variance in the dependent variable can actually be explained by taking into consideration the autoregressive term, while the second highest contribution comes from the introduction of the contemporaneous rate of inflation, which becomes more important for the upper tail.

3.4.2 What Are the Determinants of Changes in Inflation Expectations?

In order to capture the determinants of monthly changes in inflation expectations, the following percentile time series regression has been specified:

[22] As it is well known, 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.

$$\begin{aligned}
\pi^k_{t|t-12} - \pi^k_{t-13|t-25} &= \alpha + \sum_i \beta_i \left(\pi_{t-i} - \pi^k_{t-i|t-i-12}\right) \\
&\quad + \sum_j \psi_j \left(\pi^k_{t-j|t-j-12} - \pi^k_{t-j-12|t-j-24}\right) + \boldsymbol{\gamma}\boldsymbol{\Delta}\mathbf{X}_t + \varepsilon_t; \quad (29) \\
k &= 1, ..., 99; \qquad i = 13, 14; \qquad j = 1, 2; \\
\mathbf{X}_t &= \begin{bmatrix} y_t & i_t & r_t & i_{t-1} - r_{t-1} & \pi^F_{t|t-12} \end{bmatrix}',
\end{aligned}$$

where the operator Δ denotes the difference between the current value of the variable and its lagged (13 periods backwards) counterpart.

Results The second model aims to explaining what determines changes in the forecasts. In order to address this issue we select as regressors the first difference of the variables introduced in the previous model. It turns out that the explanatory power of the regressors is quite poor, apart from the first autoregressive term which has a high partial contribution coefficient along the whole distribution, as can be seen in Figure 33 (see also figures in the Appendix B). Thus, we are switching off the effect channeled by interest rate variables. Its contribution starts decreasing only from the 70^{th} percentile, leaving room for the last observed error and the second autoregressive term. The overall coefficient of determination still displays a hump-shaped pattern in the middle range. This model can actually be treated as extended model of the estimated simple learning rules. As in that model, here we can also observe the coefficient on the observed past forecast error to be significant on the RHS of the distribution.

Insert Table 4 and Figure 33 about here

3.4.3 What Are the Determinants of Errors in Inflation Forecasts?

To investigate more in depth the nature of the forecast error we estimate the following relations: evidence of serial correlation in the forecast error process indicates that there is an inefficient exploitation of information from last year's forecast in generating current year's forecast, hence violating the rationality hypothesis. Furthermore, in order to capture the possibility of an efficient exploitation of relevant information, we include in the set of regressors the SPF forecast error

$$\begin{aligned}
\pi_t - \pi^k_{t|t-12} &= \alpha + \beta\left(\pi_{t-13} - \pi^k_{t-13|t-25}\right) + \delta(\pi_t - \pi^F_{t|t-12}) + \boldsymbol{\gamma}\boldsymbol{\Delta}\mathbf{X}_t + \varepsilon_t \quad (30) \\
k &= 1, ..., 99; \qquad \mathbf{X}_t = \begin{bmatrix} y_t & \pi_t & i_t - r_t \end{bmatrix}'
\end{aligned}$$

The latter regression is similar to the Panel-D regression performed in Mankiw, Reis and Wolfers (2003) and Ball and Croushore (1995), but it is designed in a slightly different way, as it has errors and changes of relevant variables as dependent variables while the mentioned papers adopt dependent variables similar to those introduced in the our first percentile regression model.

Results The resulting response functions are reported in various figures in Appendix B. It turns out that the coefficient associated to horizontal spread is never found to be significantly different from zero, at any percentile. The same evidence holds for the coefficient of the cyclical component, but just from the 45^{th} percentile onwards, while in the previous range it has a negative sign. It is also worth noting that the function built up with a constant is downward sloping and crosses the zero line in correspondence of the 51^{th} percentile, which is classically associated with the "rational" group. The response function associated to the last observed forecast error is fairly constant up to the 30^{th} percentile and it assumes a marked U-shaped pattern afterwards. As regards the average error of the professional forecasters, which on theoretical grounds is actually expected to get a significant and positive coefficient, we can actually see that the response is first constant and then hump shaped around 55^{th} percentile, while it decreases in the last deciles.

Insert Figure 34 about here
Insert Table 5 about here

The most important inference probably comes from the observation of the coefficient of determination and from the partial "contribution" coefficients associated to each regressor. The first one declines as we move towards the upper end of the distribution, but not monotonically, displaying a quite marked hump-shaped pattern in the first two ranges and assuming a U-shaped pattern from the 70^{th} percentile. This evidence has important implications for the informational structure underlying each group. The interpretation will be more clear cut after observing the partial contribution coefficients. It appears that the last observed error has a great importance for the first range, which usually displays a backward looking "adaptive" behaviour. This might be due to agents not observing current inflation as their error could be explained by the past errors, i.e. they are just making inflation expectations around focal points such as 0 or 5 percent (digit preference). The variance of forecast errors of the third group, located in the upper end is almost exclusively explained by the variance of the change in the actual inflation. Total R^2 decreases, implying that agents are observing all the variables, although their error could not be explained by them, but just by the constant term. This is a further signal of the "pessimism" characterising these agents. But there is another possible interpretation arising from these results, as the change in the inflation is the most important variable and the rise in inflation decreases forecast errors. Together with the fact that the autoregressive component has almost no explanatory power we are lead to conclude that agents comprised in this part of the distribution might be behaving in line with what suggested by recent literature on inattentiveness and rationally heterogeneous expectations. We can notice that in the middle range the contribution of the past error decreases, while the contribution of the error of the professional forecasters gains further importance. Considering the professional forecasters as a "general" stereotype of rational agents, we can actually infer that the middle range, especially around the $50^{th} - 55^{th}$ percentile, is the least biased, as the evidence arising from the test of unbiasedness in the pervious section. In that region the error of professional forecasters is actually almost the only important variable for the determination of the forecast error. This equation could be considered as a test of rationality. The test could be that the $\alpha = \beta = \gamma = 0$. The only significant could be δ. We have tried to add several lags of the SPF to the equation to assess the Carroll's (2003a,b) finding that the transmission effect from professional forecasters to households is quite slow, but in our case additional lags tend to have no explanatory power.

4 Discussion

The evidence arising from the analysis in the previous section generally confirms the presence of a marked degree of heterogeneity in the process of expectation formation. Relying on a visual impression obtained from the models presented, we can identify (at least) three intervals of marginal response of the dependent variable to the regressors introduced in the estimation. This evidence might be due to the existence of different models of expectation formation for the individuals comprised in the overall distribution.

On empirical grounds, we can roughly consider the first interval, the one at the poor hand of the distribution, as the one characterised by agents that do not observe (or do not take into account) the relevant variables for producing one-year-ahead inflation expectations. On the other hand, individuals in the interval corresponding to the upper tail, although observing the relevant information, seem to overreact to movements in the regressors, implying a high degree of "pessimism". Intuitively, the middle range of response should comprise rational individuals.

We start with the analysis of the LHS interval of the distribution. In the Figure 35 we can observe the SSE of competing models for this part of the distribution. As we pointed out before agents comprised in this interval are behaving highly autoregressively. Nevertheless, by comparing the estimated models, it is possible to notice some asymmetries among agents in this range. Roughly speaking we can divide this interval into a further three intervals.

In this first interval (up to 10^{th} percentile) agents have nearly static expectations as they virtually never update their information sets. These agents do not take into account past inflation when forecasting inflation, but just their past forecasts and possibly to some extent the cycle indicator. Also some support of updating their parameters (on past forecast) with respect to the past errors has been found in the previous section. We can conclude that agents in this group are mainly using some form of AR(1) rule with their past forecasts. They are updating from time to time their coefficients with respect to the last observed error.

Insert Figure 35

The second interval on the LHS is comprised of agents roughly between 11^{th} and 30^{th} percentile. Also in this interval agents are not updating their information sets regularly, but they are slowly starting to implement past inflation in their PLMs and thus they are slowly moving away from static expectations. So their forecast errors are slightly less associated with their past forecast errors than in the previous group, although this dependence is only slowly decreasing. We could featured this group as a transitional group. No form of adaptive behaviour is significant for this group, except the tests for perpetual learning and convergence, which are implying that all agents are learning. We could characterise the behaviour of this group by PLM with intercept, their past forecasts and past inflation.

The third interval incorporate agents between 31^{th} and 49^{th} percentile. These agents are now starting to more regularly update their information sets, especially after the 40^{th} percentile. They fully employ past inflation into their PLMs and also information from SPF. Although we would expect that these agents are at least partly behaving in adaptive manner, no such behaviour was found in the previous section, except the Kalman filter learning which gives optimal gains in each period. In this interval the dependence of forecast errors on their past errors is further decreasing and by the end of this interval is virtually null. PLM that best characterise the behaviour of agents comprised in this interval consists of intercept, their past forecasts, past inflation and forecast of professional forecasters.

The group of agents in the middle of the empirical distribution is comprised of rational agents. Roughly speaking agents between 50^{th} and 55^{th} percentile are rational. They are updating their information sets regularly and they do not make systematic errors when forecasting inflation. The error of SPF is the only explanatory variable of their errors. As they are updating information sets regularly they do not have to rely on any form of adaptive behaviour and only some agents at the upper bound of the interval are updating their coefficient estimates with respect to new information that becomes available in the economy.

The RHS of the distribution is comprised by agents who are behaving in accordance with theories of adaptive learning and rational inattention. Figures 36 and 37 are plotting the SSE of competing models for this part of the distribution. Agents above 56^{th} percentile can be further divided into four groups. The first group is comprised of agents roughly between 56^{th} and 66^{th} percentile. The predominant feature of these agents is (dynamic) sticky information. These agents are also associated with adaptive behaviour, especially they are updating their coefficients with respect to new information. Agents in this group are on average updating their information sets every 8 to 30 months. The "pessimism" of agents in the RHS of empirical distribution is starting to show in this group as the forecast errors are more and more associated with changes in inflation.

The second group on the RHS of the distribution encompass agents between 67^{th} and 72^{rd} percentile. Agents in this group are mainly identified with the third case adaptive learning, i.e. updating coefficients with respect to new information. Moreover, we can argue that decreasing gain learning when agents have in their PLMs inflation and SPF is the closest explanation of their behaviour. Similarly to the previous interval on the RHS of the empirical distribution also agents in this group are associated with dynamic version sticky information and thus to the rational inattention argument.

Insert Figures 36 and 37

Agents between 73^{rd} and 90^{th} percentile incorporate the third group on the RHS of the empirical distribution. Their behaviour is best explained with the constant gain version of the adaptive learning where agents only observe their past forecasts and they update their coefficients with respect to the last observed error. Some support has also been found for other versions of adaptive learning. Agents in this group are updating information less frequently than agents in previous groups and the change of inflation is becoming the main determinant of their forecast errors.

The fourth group of agents on the RHS of the empirical distribution (above 91^{st} percentile) is again more associated with decreasing gain version of updating coefficients with respect to new information, although the fit of the model is considerably worse than for pervious groups of agents. Agents in this group are updating their information sets very infrequently and the change in inflations is almost the exclusive determinant of forecast errors.

We also found some support for time-varying degree of heterogeneity among agents. We were analysing that by splitting the sample into two subsamples and analysing the heterogeneity in expectations for the two different inflation regimes. As Branch (2004, 2005) found some support for time varying degrees of inflation heterogeneity we also found some support for that, although we cannot analyse it beyond splitting the sample into two subsamples. Nevertheless, the results suggest that agents slightly below the median have become more rational in the second subsample and that they are updating information more frequently[23]. Agents slightly above the median are producing systematic errors in the second subsample, but not in the first subsample.

[23] On the other hand one could interpret these results as a critique to our approach which does not allow for "switching" of agents between different percentiles. Our approach can not accomodate time varying degrees of heterogeneity, but only reports the average heterogeneity accross the sample.

Concluding Remarks

Assuming some sort of bounded rationality has become "popular" lately in macroeconomics. Especially studies of optimal monetrary policy have focus on bounded rationality and found significantly different policy recomendations if assuming that agents are boundedly rational. This is just one rationale why it is necessary to further explore how agents are forcasting key macroeconomic variables. In this study we confirm the presence of a marked degree of heterogeneity in the process of expectation formation characterising households comprised in the Michigan Survey. Our approach, aimed at identifying the main sources and degree of heterogeneity, has been pursued under different perspectives suggested by modern literature on learning, heterogeneous expectations and informational stickyness. Furthermore, we have investigated the dynamics of higher moments of the distribution of beliefs providing new evidence on their cyclical pattern and their behaviour under different inflationary regimes. As a matter of fact, and not surprisingly, we find that the variance is counter-cyclical. An interesting finding is that skewness and kurtosis are pro-cyclical, both decreasing in recessions and increasing in expansionary periods. Also in the period of stable inflation the variance is less volatile, while skewness and kurtosis are more volatile: this has important implications under a rational inattentiveness story.

We provide among the first contributions of learning under heterogeneous expectations, by identifying at least three different subgroups of the overall population of respondents, each of them characterised by different informational sets and different learning mechanisms: these groups can be divided into further sub-components regarding their informational sets and their learning dynamics, if any has been detected.

Tests for "static" and "dynamic" versions of sticky information have also been conducted. Only agents in the middle of the distribution are regularly updating their information sets. Evidence of rational inattention has been found for agents comprised in the upper end of the distribution. We identify three regions of the overall distribution corresponding to different expectation formation processes, which display a heterogeneous response to main macroeconomic indicators: a static or highly autoregressive (LHS) group, a "nearly" rational group (middle), and a group of agents (RHS) behaving in accordance to adaptive learning and sticky information. The latter are generally responding in a too "pessimistic" way, as they tend to overreact to macroeconomic fluctuations. Our dynamic version of the test for sticky information suggests that agents are more likely to regularly update their information sets when inflation "matters", especially for the center-right hand side of the distribution.

The adoption of different tests for learning suggests that agents on the right hand side (RHS) of the distribution tend to behave in an adaptive manner, whereas agents on the left hand side (LHS) of distribution do not exhibit such behaviour. Agents on the RHS of the distribution are particularly associated with updating their coefficients with respect to "new information" and slightly less to updating their coefficients with respect to past errors. To further investigate this issue, we estimate several additional time series models of expectation formation. These models confirm a significant degree of heterogeneity and asymmetry in the expectation formation process.

As a proposal for further study, it would be interesting to deepen the analysis on learning behaviour by allowing agents to endogenously switch between different algorithms, especially after structural breaks occur. We conjecture that this hybrid learning mechanism would probably lead to an even better fit of the data. Furthermore, we suggest that some combination of adaptive learning and rational inattention would probably reach a better descriptive performance for agents in the RHS of the empirical distribution.

References

[1] Baak, S.-J. (1999). 'Tests for Bounded Rationality with a Linear Dynamic Model Distorted by Heterogeneous Expectations', Journal of Economic Dynamics and Control, vol. 23(9-10), (September), pp 1517-43.

[2] Bakhshi, H. and A. Yates (1998). 'Are UK Inflation Expectations Rational?', Bank of England Working Paper No. 81.

[3] Ball, L. and D. Croushore (1995). 'Expectations and the Effects of Monetary Policy'. NBER Working Paper 5344.

[4] Basdevant, O. (2003). 'Learning Process and Rational Expectations: An Analysis Using a Small Macroeconomic model for New Zealand' Reserve Bank of New Zealand Discussion Paper Series DP2003/05.

[5] Branch, W. A. (2004). 'The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations', Economic Journal, vol. 114, (July), pp 592-621.

[6] Branch, W. A. (2005). 'Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations'. Mimeo. University of California, Irvine.

[7] Brock, W. A. and C. H. Hommes (1997). 'A Rational Route to Randomness', Econometrica, vol. 65(5), (September), pp 1059-95.

[8] Bryan, M. F. and S. Palmqvist (2005). 'Testing Near-Rationality Using Detailed Survey Data' Federal Reserve Bank of Cleveland working paper 05-02.

[9] Carceles-Poveda, E. and C. Giannitsarou (2005), 'Adaptive Learning in Practise', Mimeo. University of Cambridge.

[10] Carroll, C. D. (2003a). 'Macroeconomic Expectations of Households and Professional Forecasters', Quarterly Journal of Economics, vol. 118(1), (February), pp 269-98.

[11] Carroll, C. D. (2003b). 'The Epidemiology of Macroeconomic Expectations', Mimeo. Johns Hopkins University.

[12] Chavas, J.-P. (2000). 'On Information and Market Dynamics: The Case of the U.S. Beef Market', Journal of Economic Dynamics and Control, vol. 24(5-7), (June), pp 833-53.

[13] Croushore, D. (1998). 'Evaluating Inflation Forecasts', Federal Reserve Bank of Philadelphia Working Paper number 98-14.

[14] Curtin, R. (1996). 'Procedure to Estimate Price Expectations'. Mimeo. University of Michigan.

[15] Curtin, R. (2005). 'Inflation Expectations: Theoretical Models and Empirical Tests'. Mimeo. University of Michigan.

[16] Döpke, J., J. Dovern, U. Fritsche, and J Slacalek (2006). "The Dynamics of European Inflation Expectations," DIW Berlin Discussion Paper No. 571.

[17] Ezekiel, M. (1938). 'The Cobweb Theorem', Quarterly Journal of Economics, vol. 52(2), (February), pp 255-280.

[18] Evans, G. W. and S. Honkapohja (2001). 'Learning and Expectations in Macroeconomics', Princeton University Press. Princeton and Oxford.

[19] Evans, G. W. , S. Honkapohja and N. Williams (2005). 'Generalized Stochastic Gradient Learning'. NBER Technical Working Paper No. 317

[20] Fischer, I. (1930). 'Theory of Interest'. Macmillan, New York.

[21] Fuchs, G. (1979). 'Is Error Learning Behaviour Stabilizing?' Journal of Economic Theory, vol. 20, pp 300-17.

[22] Granger, C. W. J., and T. Teräsvirta (1993). 'Modelling Nonlinear Economic Relationships'. Oxford University Press, New York.

[23] Hall, S. J., D. Robertson and M. R. Wickens (1997). 'Measuring Economic Convergence', International Journal of Finance and Economics, vol. 2(2), (April), pp 131-43.

[24] Jonung, L. (1981). 'Perceived and Expected Rates of Inflation in Sweden', American Economic Review, vol. 71(5), (December), pp 961-68.

[25] Jonung, L. and D. E. Laidler (1988). "Are Perceptions of Inflation Rational? Some Evidence for Sweden," American Economic Review, American Economic Association, vol. 78(5), pages 1080-87.

[26] Kuttner K.N. (1994), Estimating Potential Output as a Latent Variable, Journal of Business and Economic Statistics, 12, 3:361-368.

[27] Llambros, L. and V. Zarnowitz (1987). 'Consensus and Uncertainty in Economic Prediction'. Journal of Political Economics, vol. 95(3), pp 591-62.

[28] Mankiw, N. G. and R. Reis (2002). 'Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve', Quarterly Journal of Economics, vol. 117(4), pp 1295-328.

[29] Mankiw, N. G., R. Reis, and J. Wolfers (2003). 'Disagreement about Inflation Expectations', NBER Working Paper No. 9796.

[30] Marcet, A. and T. Sargent. (1989). 'Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Model', Journal of Economic Theory, vol. 48, pp 337-68.

[31] Milani, F. (2005a), 'Expectations, Learning and Macroeconomic Persistence', Mimeo. University of Princeton.

[32] Milani, F. (2005b), 'Adaptive Learning and Inflation Persistence', Mimeo. University of Princeton.

[33] Muth, J.F. (1961). 'Rational Expectations and the Theory of Price Movement', Econometrica, vol. 29(3) (July), pp 315-335.

[34] Nerlove, M. L. (1958). 'Adaptive Expectations and Coweb Phenomena', Quarterly Journal of Economics, vol. 72(2), (May), pp 227-240.

[35] 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.

[36] Orphanides, A. and J. C. Williams (2003). 'Imperfect Knowledge, Inflation Expectations, and Monetary Policy', NBER Working Paper No. 9884.

[37] Orphanides, A. and J. C. Williams (2005a). 'The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations', Journal of Economic Dynamics and Control, vol. 29(11), (November), pp. 1927-50.

[38] Orphanides, A. and J. C. Williams (2005b). 'Inflation Scares and Forecast-Based Monetary Policy' Review of Economic Dynamics, vol. 8(2), pp 498-527.

[39] Pesaran, M. H. (1985). 'Formation of Inflation Expectations in British Manufacturing Industries', Economic Journal, vol. 95, pp. 948-975.

[40] Pesaran, M. H. (1987). 'The Limits to Rational Expectations' Pub. Basil Blackwell, Oxford., reprinted with corrections 1989.

[41] Pesaran, M. H. and M. Weale (2005). "Survey Expectations," Cambridge Working Papers in Economics 0536, Faculty of Economics (DAE), University of Cambridge.

[42] Pfajfar, D. (2005).'Heterogeneous Expectations in the New Keynesian Macroeconomic Model: Empirical Investigation' Mimeo. University of Cambridge.

[43] Pfajfar, D. and E. Santoro (2006).'Asymmetries in Inflation Expectation Formation Across Demographic Groups' Mimeo. University of Cambridge.

[44] Roberts, J.M. (1997). 'Is Inflation Sticky?', Journal of Monetary Economics, vol. 39(2), (July), pp 173-96.

[45] Scherrer, B. (1984). 'Biostatistique', Gaëtan Morin, Chicoutimi.

[46] Sims, C. (2003), 'Implications of Rational Inattention', Journal of Monetary Economics, vol. 50(3), (April), pp 665-90.

[47] Sims, C. (2006), 'Rational Inattention: A Research Agenda', Mimeo. University of Princeton.

5 Tables and figures

Table 1:

perc.	upsilon	c	lambda	t-test	perc.	upsilon	c	lambda	t-test
1	0.3	9	0.0094967	1.1091	50	3.69	7.4	0.15088	4.5092
2	0.35	9	0.011282	1.1682	51	3.18	7.4	0.154	4.5385
3	0.37	9	0.0089362	1.0026	52	2.02	7.5	0.15446	4.4979
4	0.37	9	0.0074702	0.89876	53	1.46	7.5	0.15086	4.3977
5	4	7.6	0.0045383	0.74861	54	1.23	7.5	0.14302	4.2431
6	4	7.6	0.0032917	0.62492	55	1.08	7.5	0.1325	4.0509
7	4	7.6	0.0025195	0.53249	56	0.96	7.4	0.12084	3.8476
8	3.56	7.2	0.0018431	0.43341	57	0.85	7.4	0.10976	3.6511
9	2.86	7.4	0.0012214	0.31442	58	0.18	1.73	0.13135	3.4807
10	0.61	3.67	-0.0010524	-0.23567	59	0.19	2.05	0.12	3.3534
11	0.58	3.67	-0.0012361	-0.26053	60	0.2	2.36	0.10937	3.2407
12	2.69	9	-0.0016425	-0.39862	61	0.21	2.56	0.099153	3.1381
13	3.72	9.2	-0.0027256	-0.6342	62	0.21	2.58	0.090365	3.0477
14	3.56	9.2	-0.0029287	-0.62028	63	0.22	2.59	0.081772	2.9686
15	3.64	9.2	-0.0030186	-0.60645	64	0.22	2.59	0.075055	2.9055
16	3.66	9.2	-0.0030574	-0.61383	65	0.23	2.63	0.069284	2.8683
17	3.64	9.2	-0.0030204	-0.61209	66	0.23	2.65	0.065379	2.8556
18	3.64	9.2	-0.0028325	-0.56024	67	0.24	2.65	0.062187	2.8671
19	3.68	9.2	-0.0023818	-0.44426	68	0.25	2.65	0.060255	2.9004
20	2.96	10.5	0.0019714	0.33185	69	0.26	2.65	0.059371	2.9498
21	2.68	10.5	0.0024806	0.38984	70	0.26	2.63	0.059715	3.0106
22	2.76	10.7	0.0031569	0.46075	71	0.27	2.6	0.060064	3.0725
23	3.3	10.7	0.0044188	0.58791	72	0.27	2.6	0.060976	3.1246
24	4.18	10.7	0.0054709	0.67675	73	0.28	2.59	0.061299	3.1602
25	5.14	10.7	0.0068537	0.78255	74	0.28	2.59	0.061771	3.1755
26	5.72	10.7	0.0079936	0.85779	75	0.28	2.56	0.061979	3.171
27	6.12	10.7	0.0087113	0.89582	76	0.26	2.05	0.061897	3.151
28	6.4	10.7	0.010278	0.99039	77	0.25	1.73	0.061641	3.1219
29	10	10.7	0.012013	1.0892	78	0.24	1.68	0.062458	3.0924
30	10	10.7	0.013535	1.1611	79	1.74	6.5	0.049372	3.1008
31	10	10.7	0.015651	1.2602	80	1.77	6.5	0.050668	3.1186
32	10	10.7	0.016573	1.2825	81	1.74	6.5	0.052161	3.1372
33	0.3	12	0.021094	1.2765	82	1.59	6.5	0.054003	3.1579
34	0.34	12	0.023859	1.4069	83	0.8	7.2	0.059801	3.1997
35	10	4.34	0.019275	1.3298	84	0.66	7.2	0.064379	3.2611
36	10	4.31	0.023337	1.5033	85	0.58	7.2	0.068168	3.3099
37	5.82	8.3	0.027745	1.6971	86	0.54	7.2	0.068044	3.2814
38	6.04	8.3	0.032936	1.8971	87	0.53	7.1	0.063286	3.157
39	6.48	8.3	0.03834	2.086	88	0.53	6.8	0.05671	2.9836
40	7	8.3	0.043773	2.2536	89	0.51	6.6	0.048995	2.7658
41	10	6.9	0.050625	2.4432	90	0.48	6.5	0.040709	2.5118
42	10	6.9	0.05921	2.6839	91	0.47	6.5	0.031581	2.2167
43	10	6.9	0.06976	2.9588	92	0.48	6.5	0.024189	1.9567
44	10	6.9	0.081709	3.245	93	0.52	6.5	0.018392	1.7343
45	10	6.9	0.094469	3.5259	94	0.78	6.5	0.013987	1.576
46	10	6.9	0.10756	3.7924	95	1.29	6.5	0.013066	1.5578
47	10	6.9	0.12059	4.039	96	8	2.81	0.014051	1.6262
48	10	6.9	0.1327	4.2526	97	8	2.82	0.014836	1.6751
49	10	6.9	0.14248	4.4121	98	8	2.82	0.013525	1.6238
50	3.69	7.4	0.15088	4.5092	99	8	2.82	0.010116	1.444

Table 2:

max. gain/lowest SSE	CGL		DGL	
SPF	$8.0 * 10^{-4}$	60.3	$3.5 * 10^{-2} t^{-1}$	53.1
OUT	$1.15 * 10^{-3}$	61.6	$5.1 * 10^{-2} t^{-1}$	61.4
AR(2)	$4.75 * 10^{-4}$	62.6	$2.5 * 10^{-2} t^{-1}$	63.3

Table 3:

Percentile	α_0	α_1	α_5	α_{11}	α_{12}	Adj R^2	DW	LM
5	-0.095	-0.011	0.105	0.661	-0.006	0.591	1.865	3.419
	-1.126	-0.446	3.864	14.812	-0.142			
	0.000	0.001	0.093	0.502	0.001			
20	0.127	0.008	0.039	0.825	-0.011	0.784	2.108	1.241
	1.954	0.382	2.238	25.629	-0.361			
	0.000	0.013	0.024	0.758	-0.008			
35	0.414	0.112	0.079	0.692	-0.055	0.811	2.086	10.782
	4.119	3.464	3.067	16.277	-1.224			
	0.000	0.208	0.030	0.631	-0.055			
50	0.644	0.142	0.053	0.642	0.029	0.884	2.090	1.600
	5.454	4.041	2.100	14.244	0.624			
	0.000	0.234	0.009	0.614	0.027			
65	0.597	0.141	0.014	0.767	-0.002	0.960	2.201	5.630
	5.679	4.617	0.725	20.205	-0.064			
	0.000	0.195	0.001	0.766	-0.002			
80	0.830	0.222	-0.018	0.575	0.267	0.926	2.170	20.817
	5.356	4.669	-0.495	12.411	3.711			
	0.000	0.213	-0.001	0.557	0.158			
95	4.923	0.310	-0.113	0.379	0.728	0.885	2.073	5.272
	10.517	4.296	-1.833	7.054	5.675			
	0.000	0.216	-0.002	0.352	0.321			

Table 4:

Percentile	β_0	β_1	β_4	β_5	β_6	β_{10}	Adj R^2	DW	LM
5	0.022	-0.006	0.025	0.050	0.818	-0.014	0.744	2.002	1.547
	0.371	-0.559	0.424	2.688	13.819	-0.336			
	0.000	0.000	0.018	0.055	0.679	-0.003			
20	0.080	-0.031	0.083	0.033	0.832	-0.031	0.857	1.959	5.297
	2.076	-3.377	1.419	2.818	14.711	-1.080			
	0.000	0.019	0.066	0.027	0.749	-0.002			
35	0.044	-0.030	0.176	0.039	0.721	0.075	0.787	1.988	0.298
	0.934	-1.791	2.973	2.302	12.517	1.529			
	0.000	-0.005	0.131	0.029	0.609	0.026			
50	-0.009	-0.016	0.210	0.032	0.733	0.073	0.819	2.037	5.649
	-0.258	-0.713	3.508	1.813	12.993	1.407			
	0.000	-0.008	0.160	0.016	0.633	0.021			
65	-0.033	-0.021	0.222	0.011	0.749	0.111	0.892	1.996	4.886
	-1.001	-0.930	3.676	0.849	13.373	2.326			
	0.000	-0.014	0.178	0.005	0.674	0.051			
80	0.435	0.181	0.211	0.029	0.553	0.088	0.794	1.906	9.318
	3.305	3.656	3.532	1.193	9.566	1.050			
	0.000	0.155	0.155	0.007	0.457	0.024			
95	2.593	0.247	0.204	0.015	0.435	0.206	0.729	2.019	3.854
	4.738	4.835	3.605	0.370	7.712	1.665			
	0.000	0.199	0.150	0.002	0.349	0.033			

Table 5:

Percentile	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	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			

Figure 1:

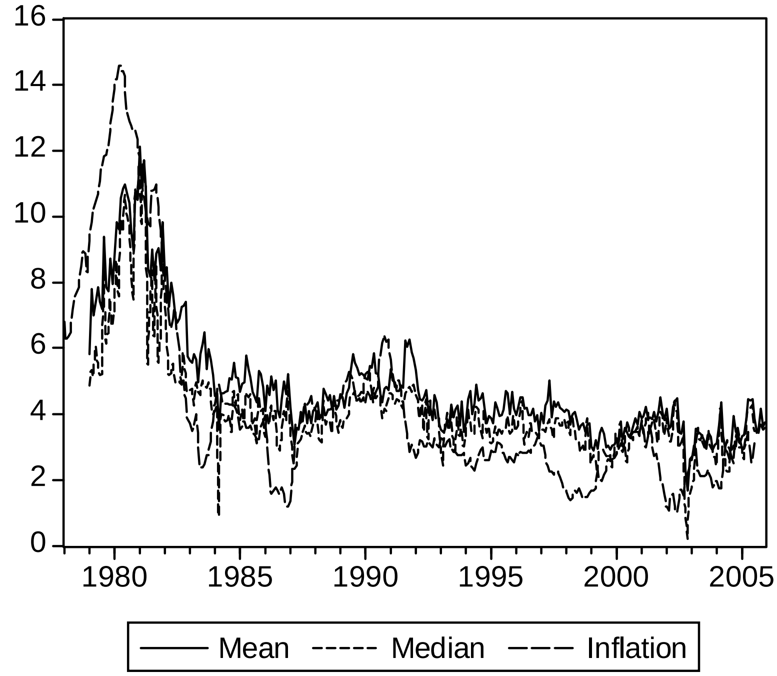


Figure 2:

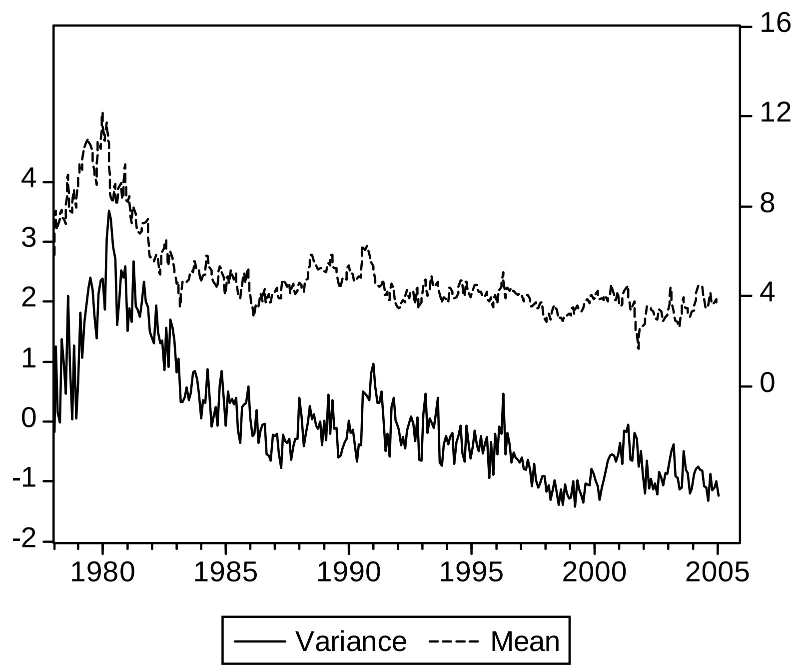


Figure 3:

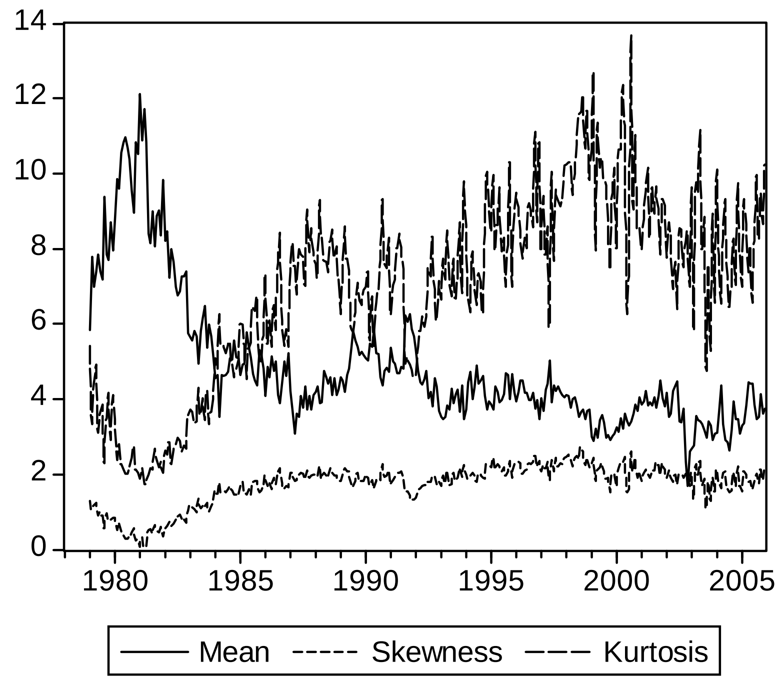


Figure 4:

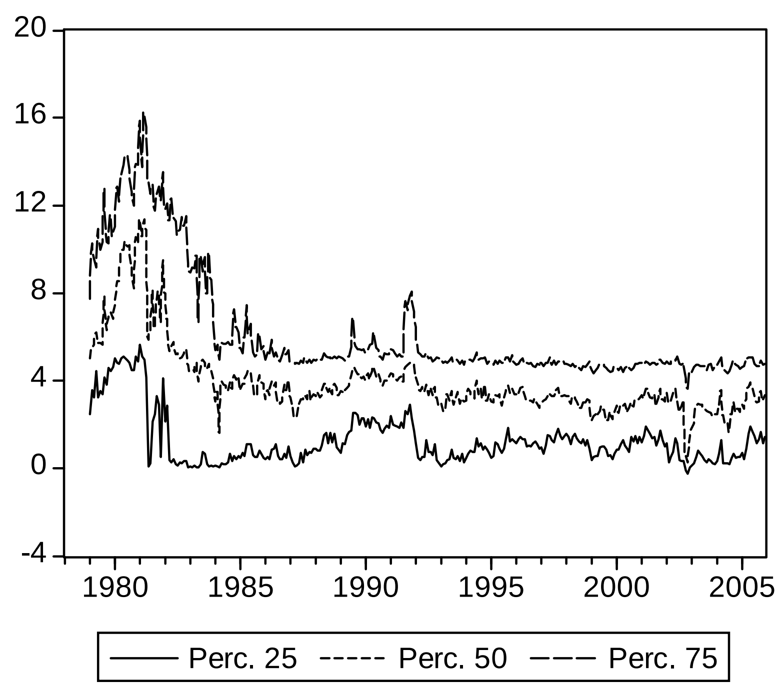


Figure 5:

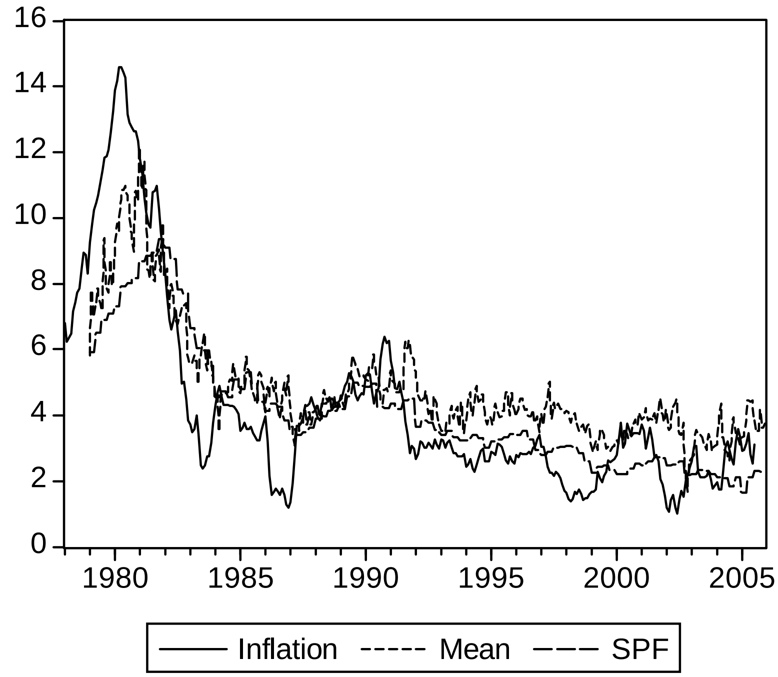
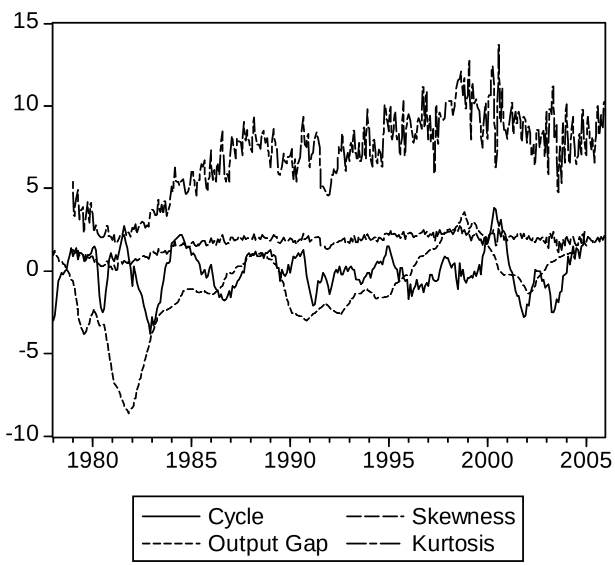


Figure 6, 7:



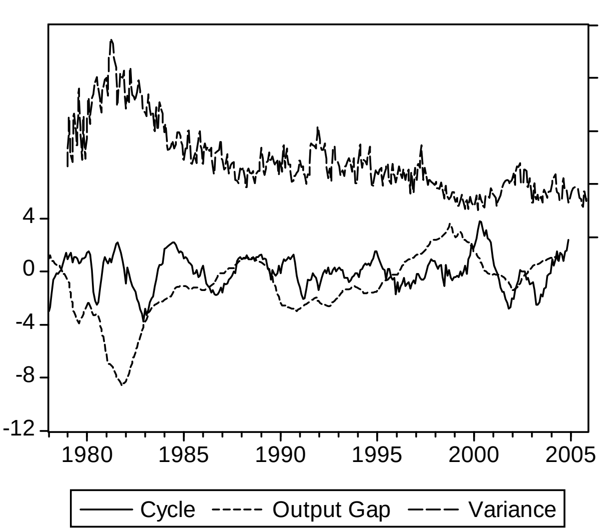


Figure 8:

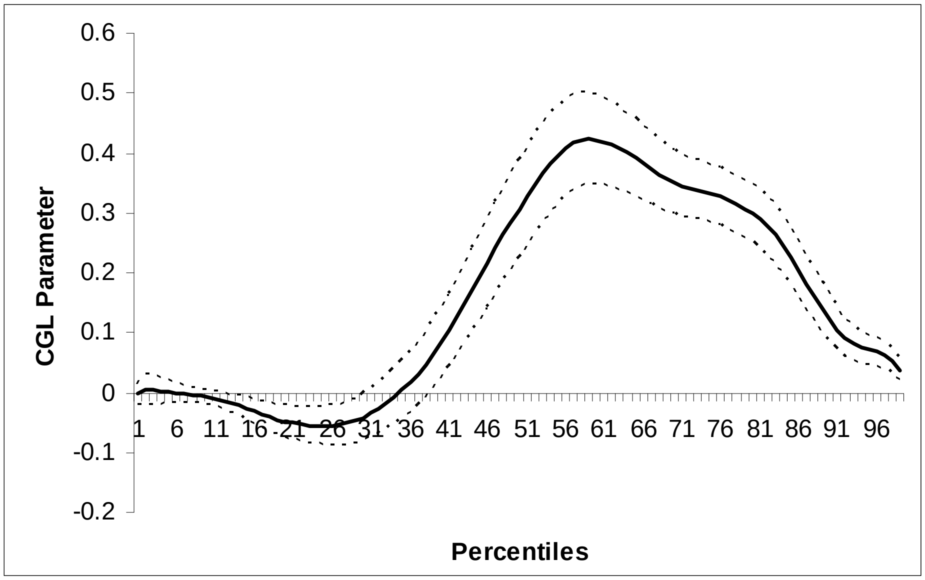


Figure 9:

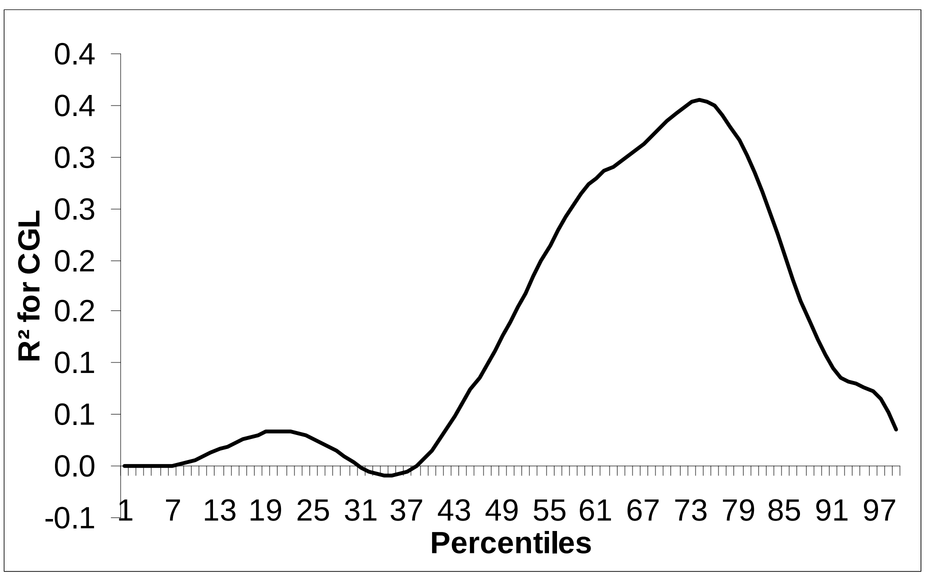


Figure 10:

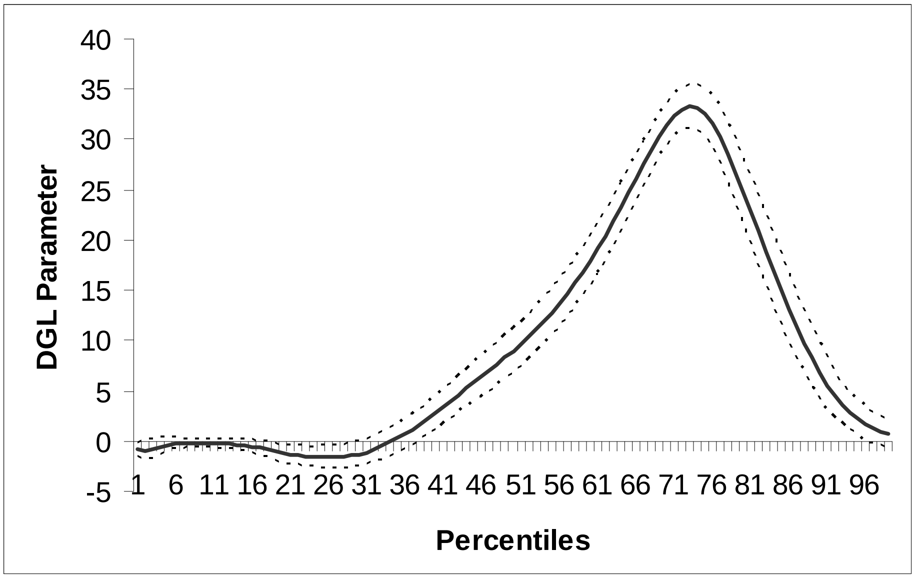


Figure 11:

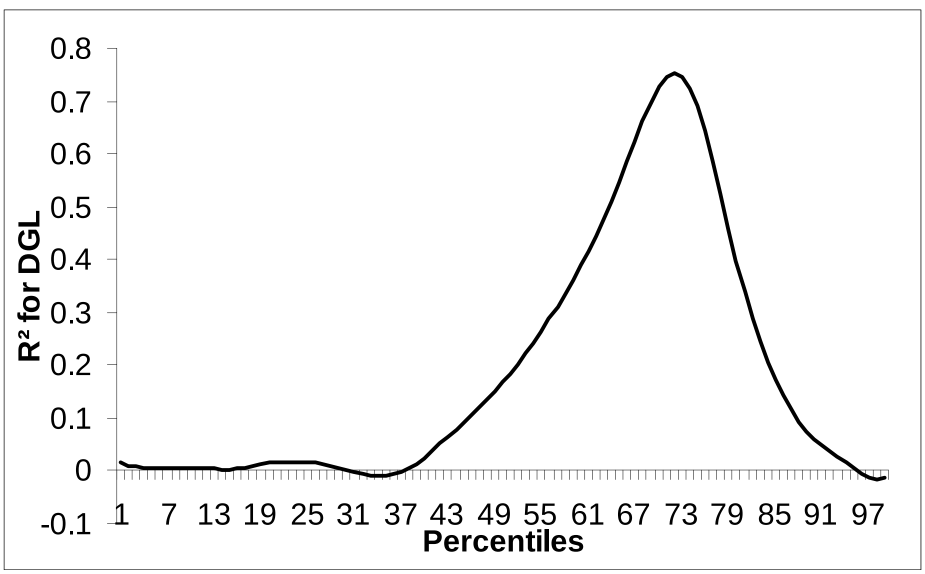


Figure 12:

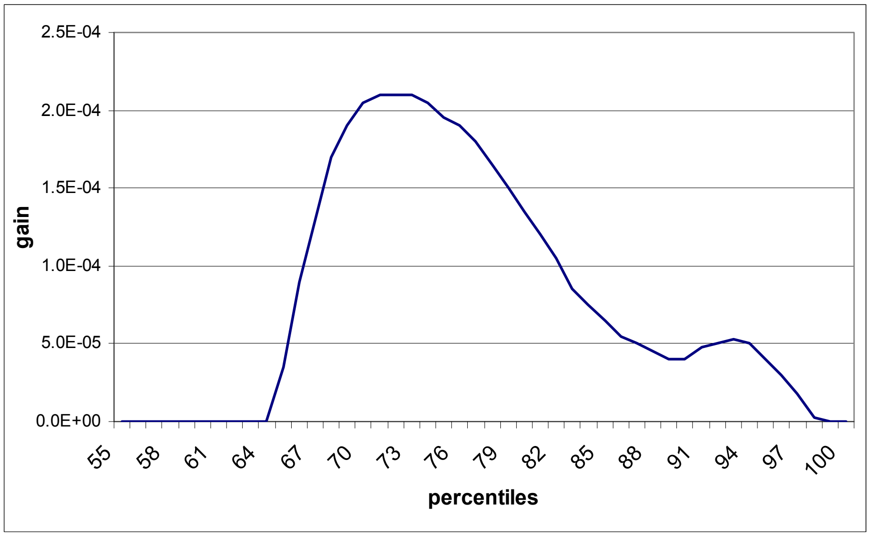


Figure 13:

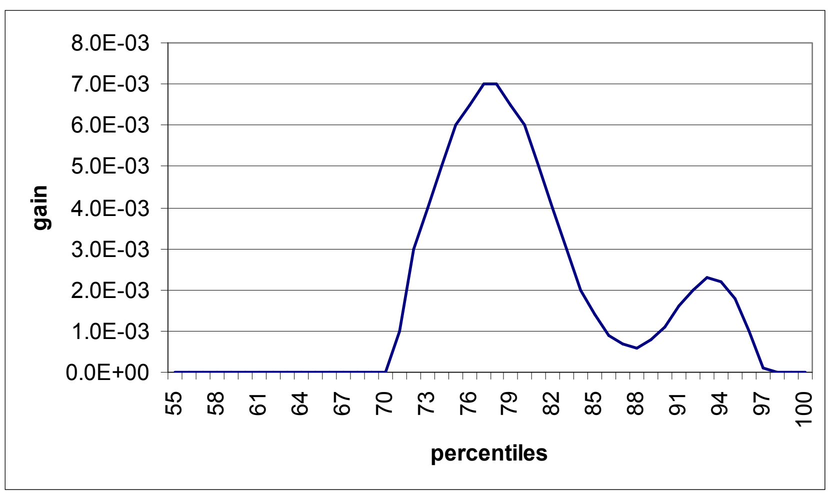


Figure 14:

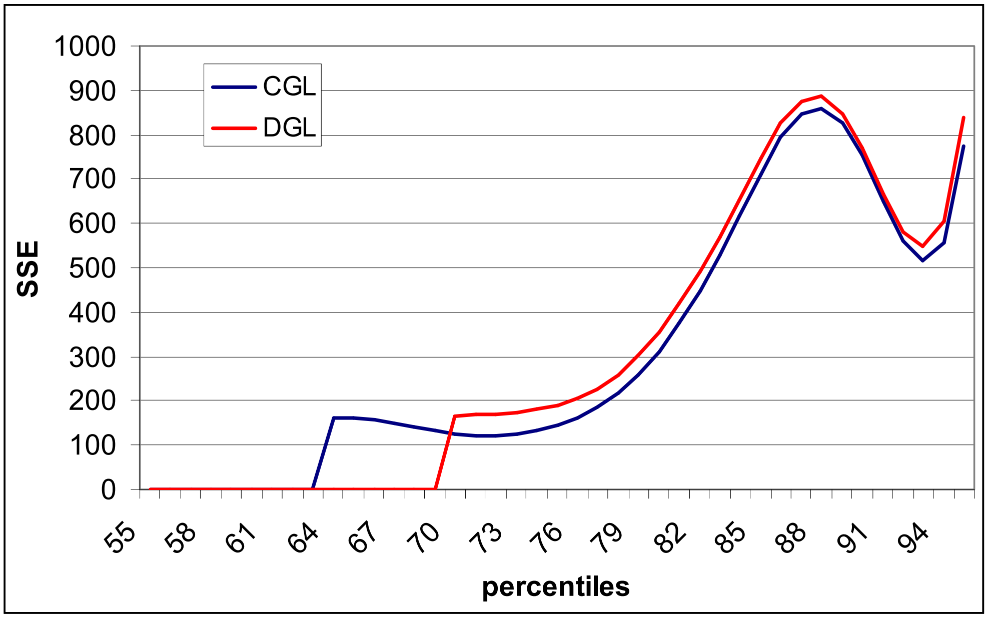


Figure 15:

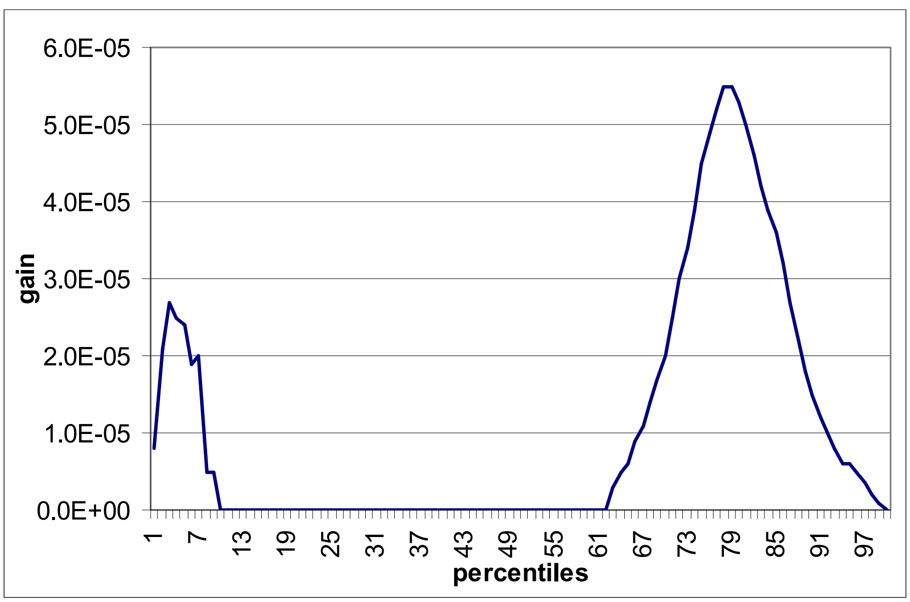


Figure 16:

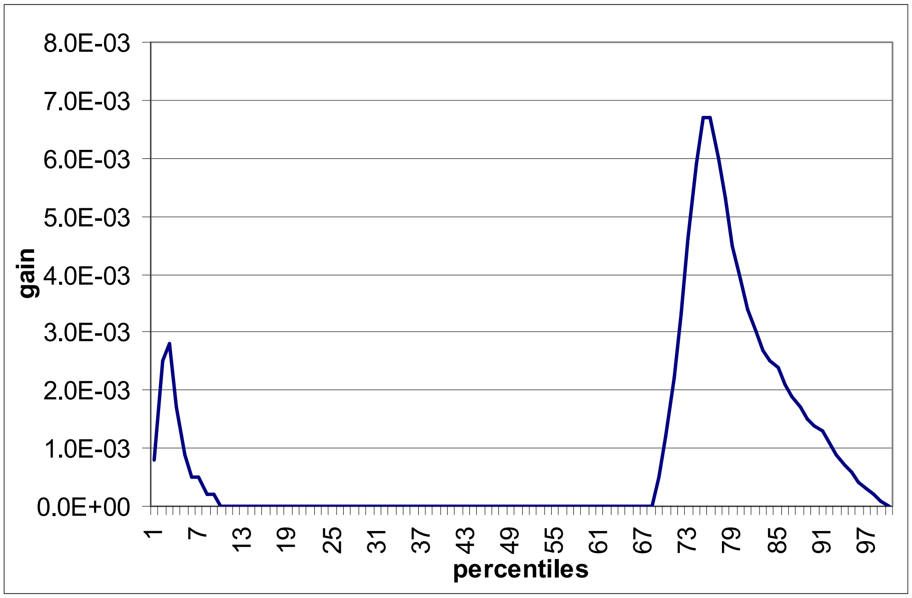


Figure 17:

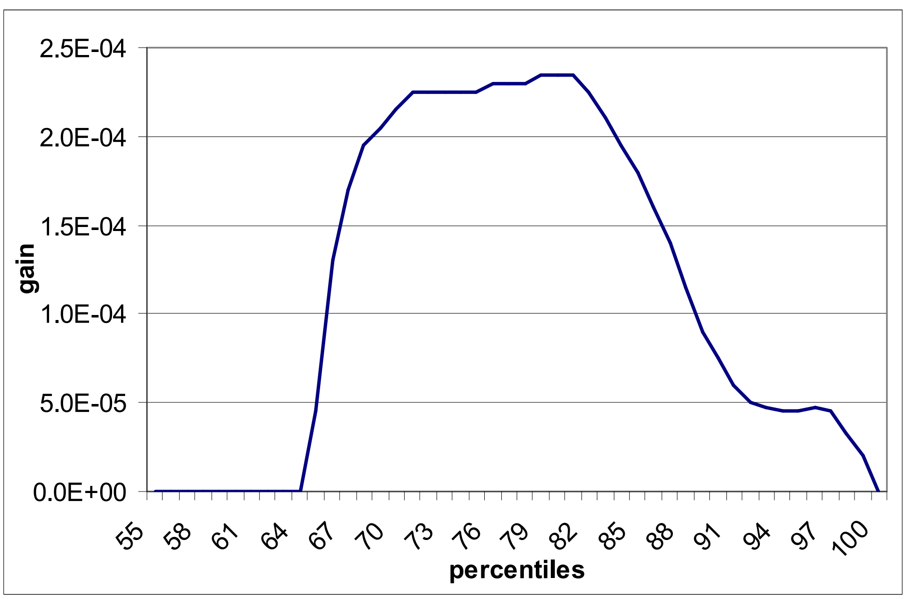


Figure 18:

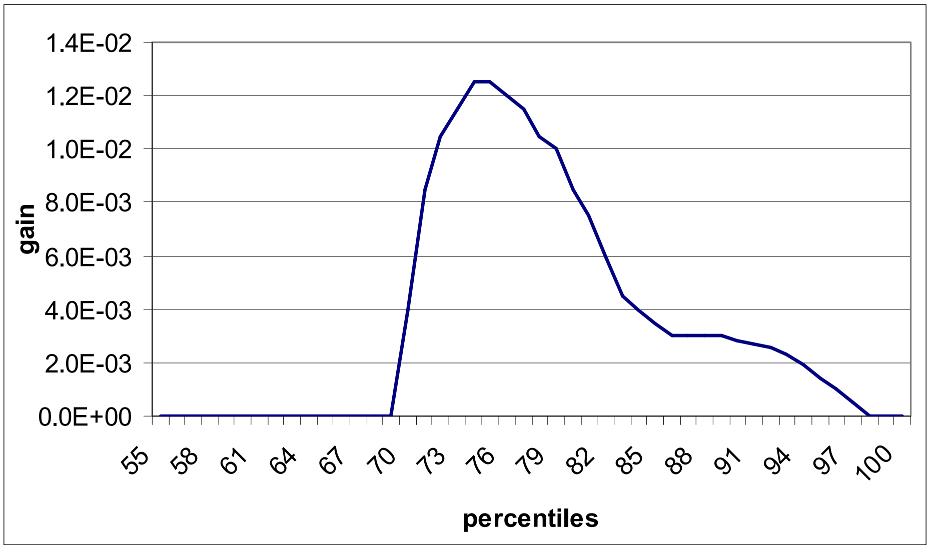


Figure 19:

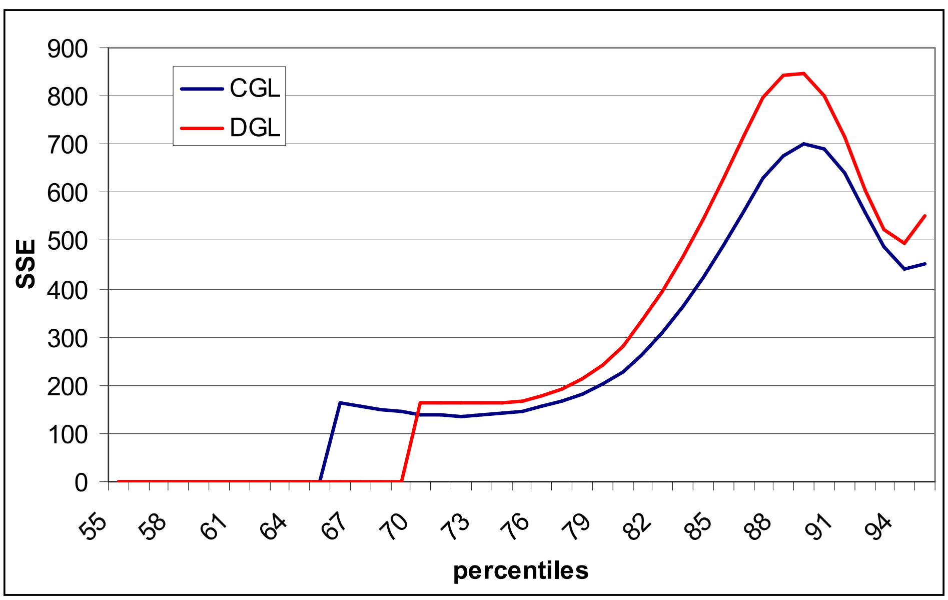


Figure 20:

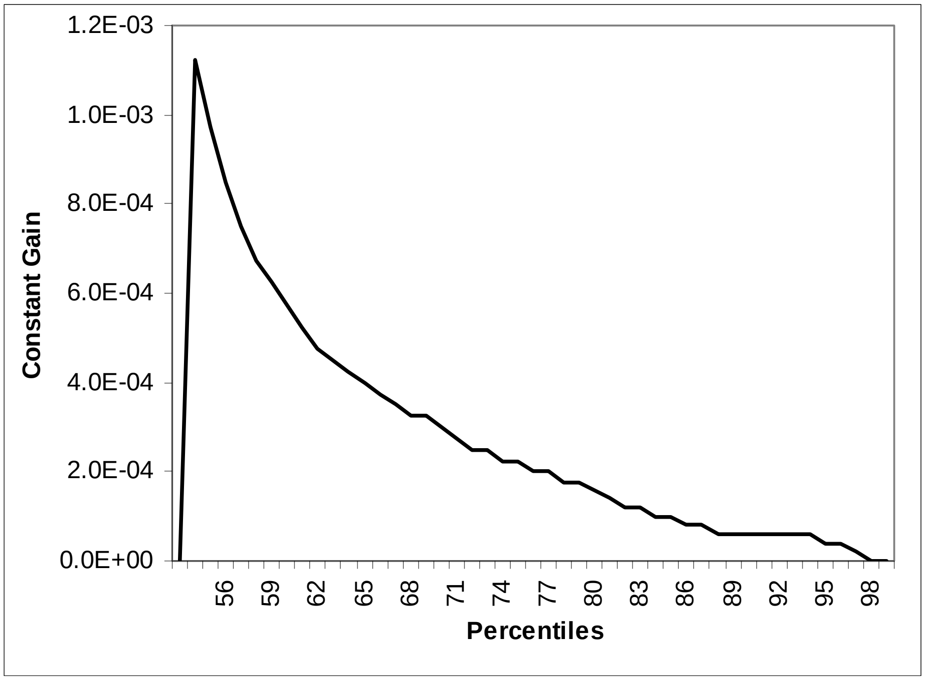


Figure 21:

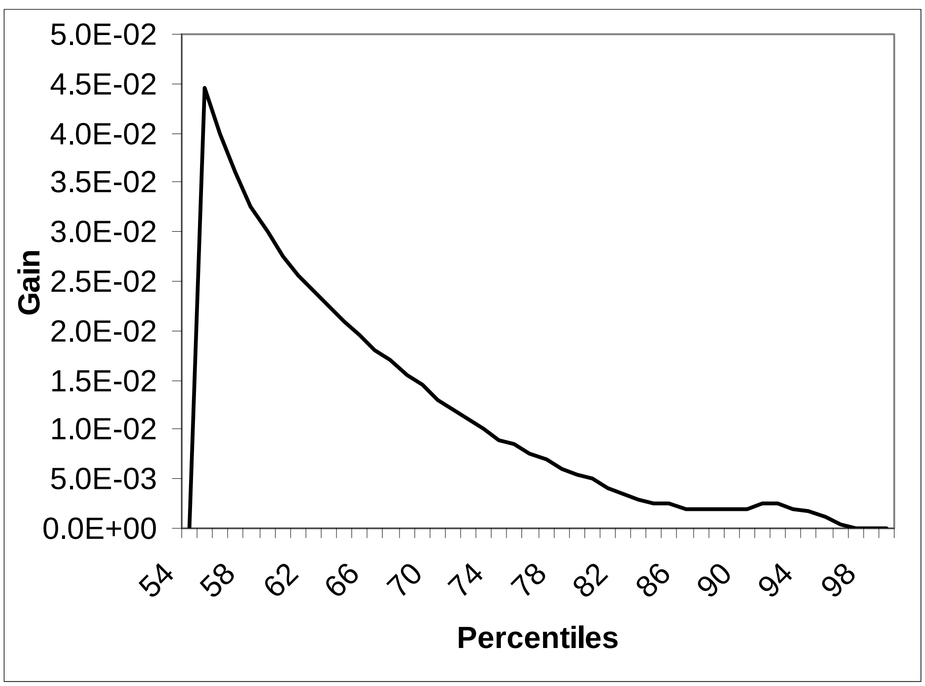


Figure 22:

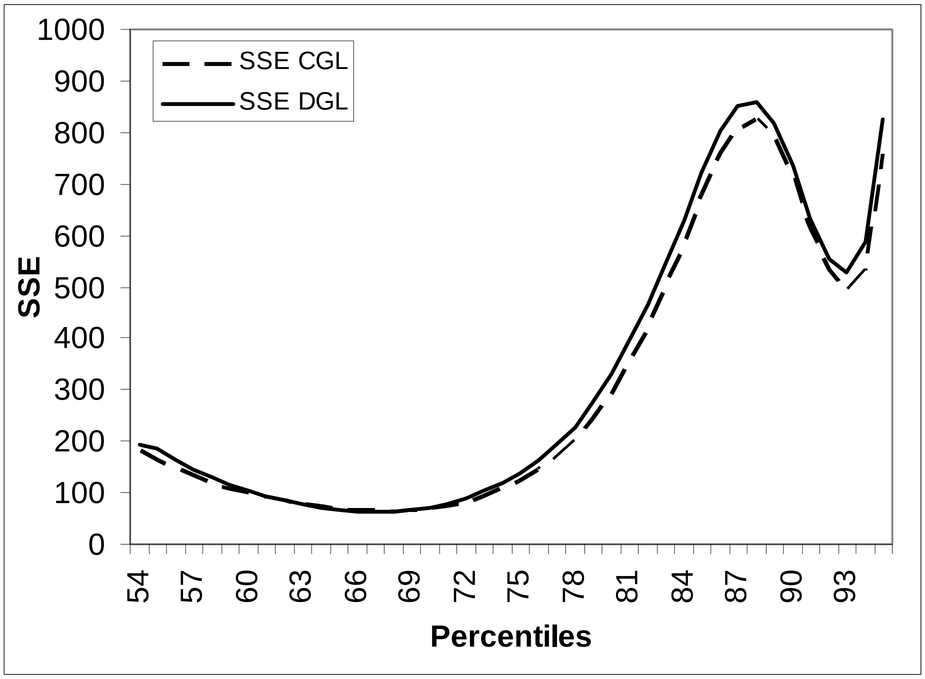


Figure 23:

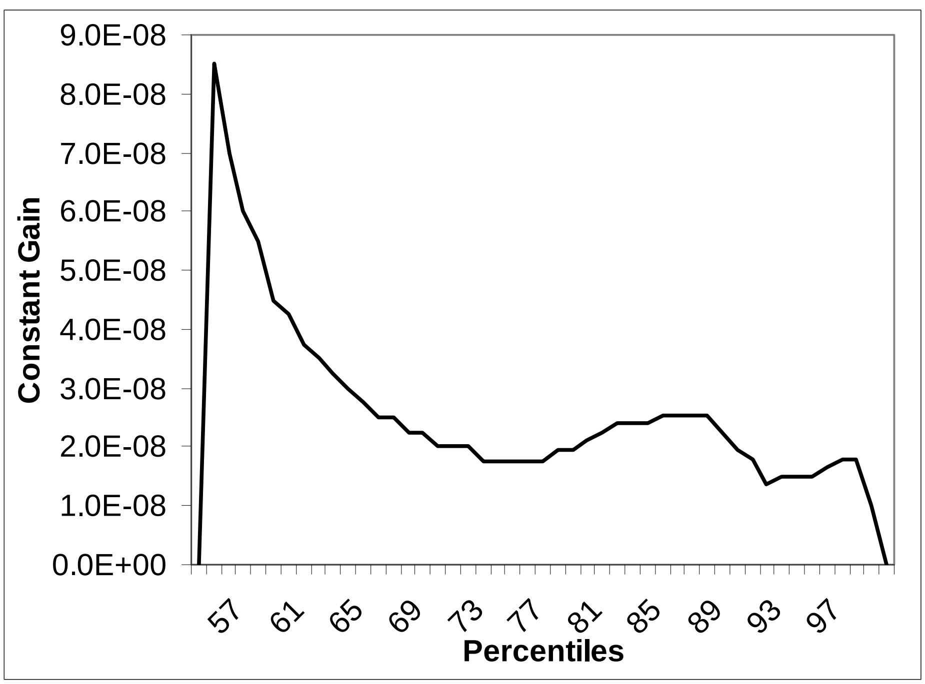


Figure 24:

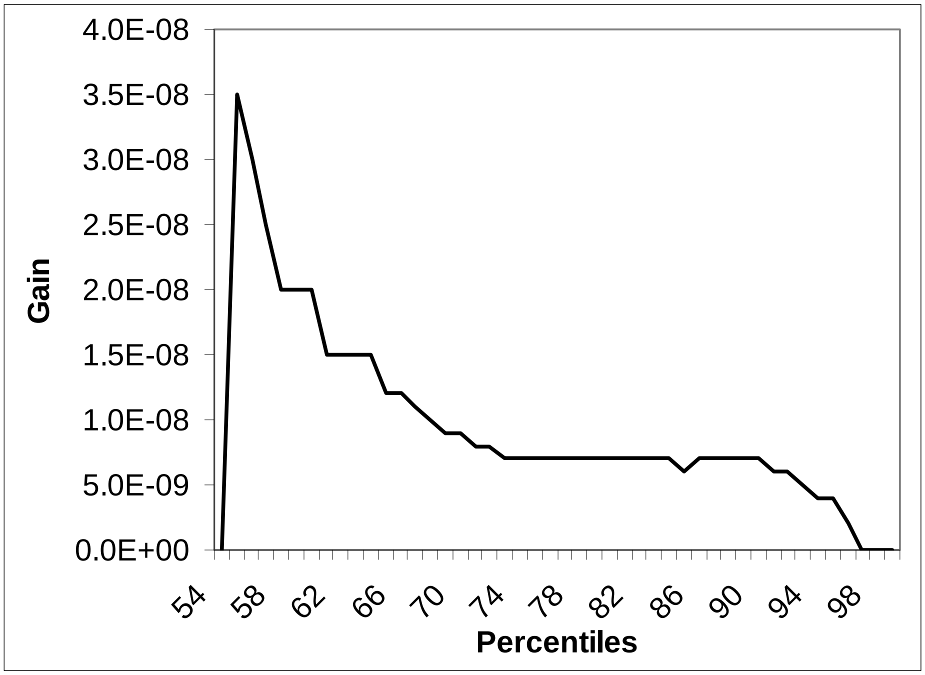


Figure 25:

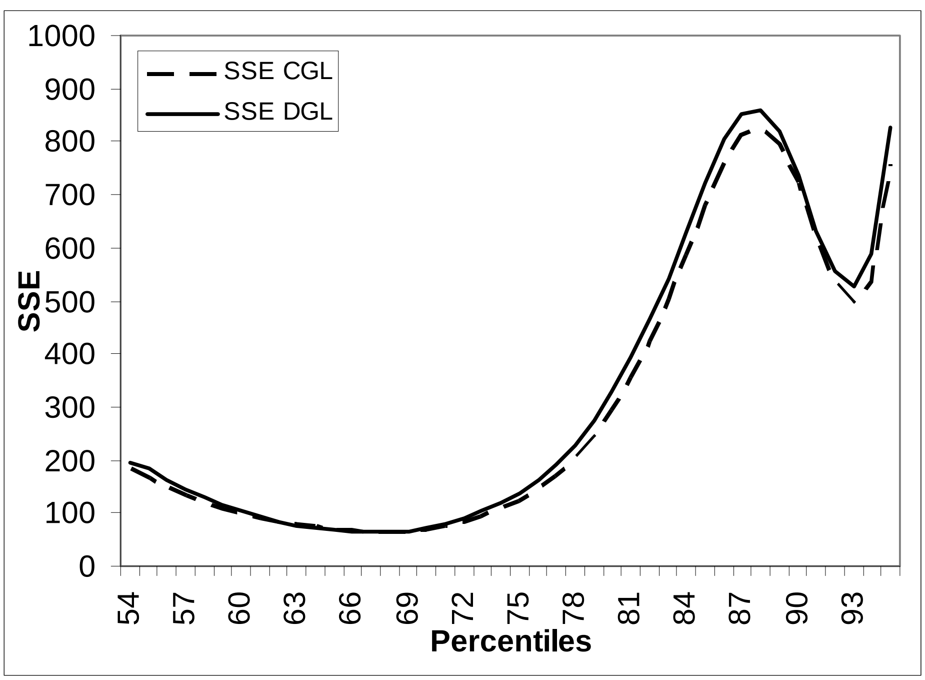


Figure 26:

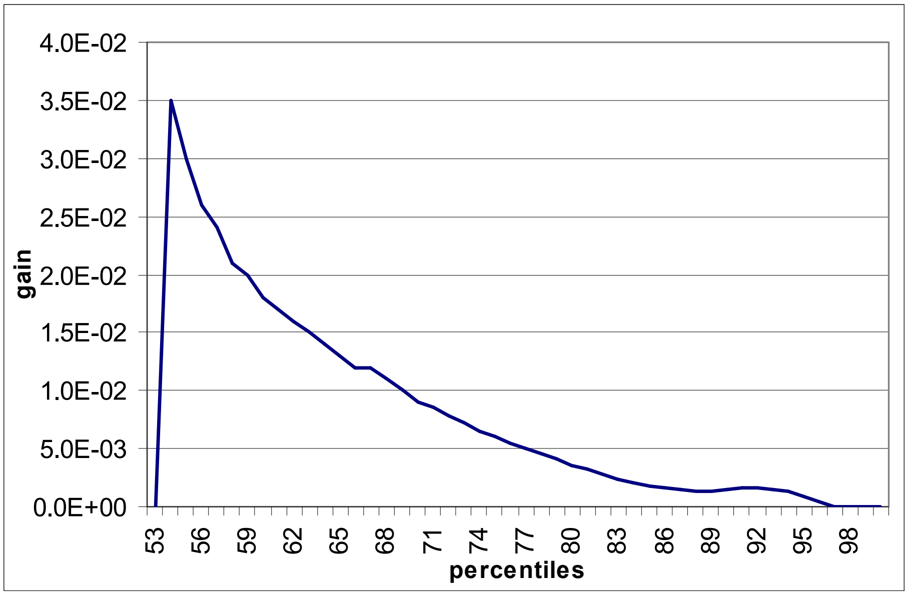


Figure 27:

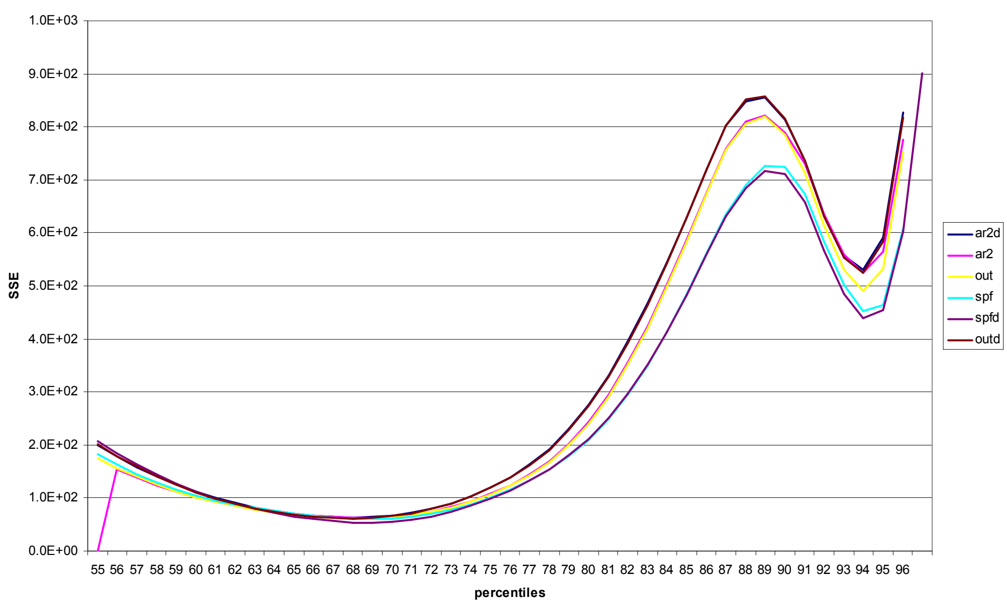


Figure 28:

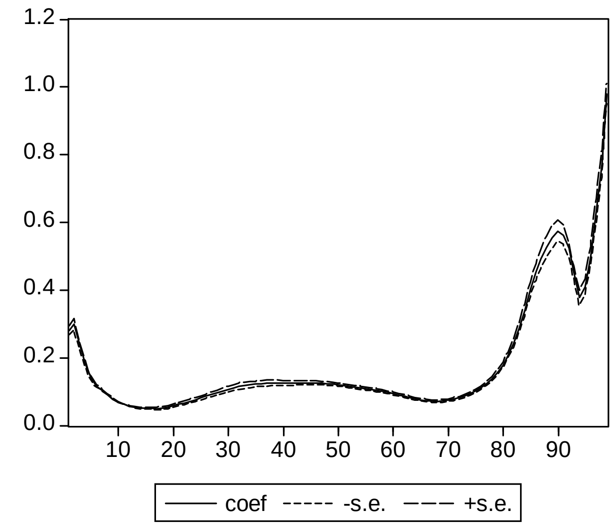


Figure 29:

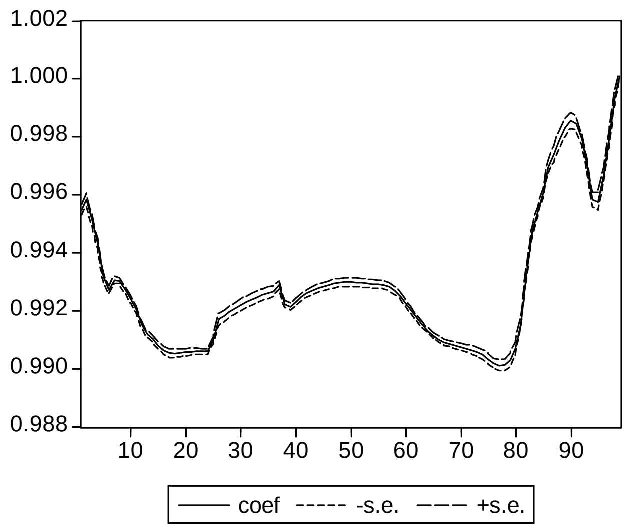


Figure 30:

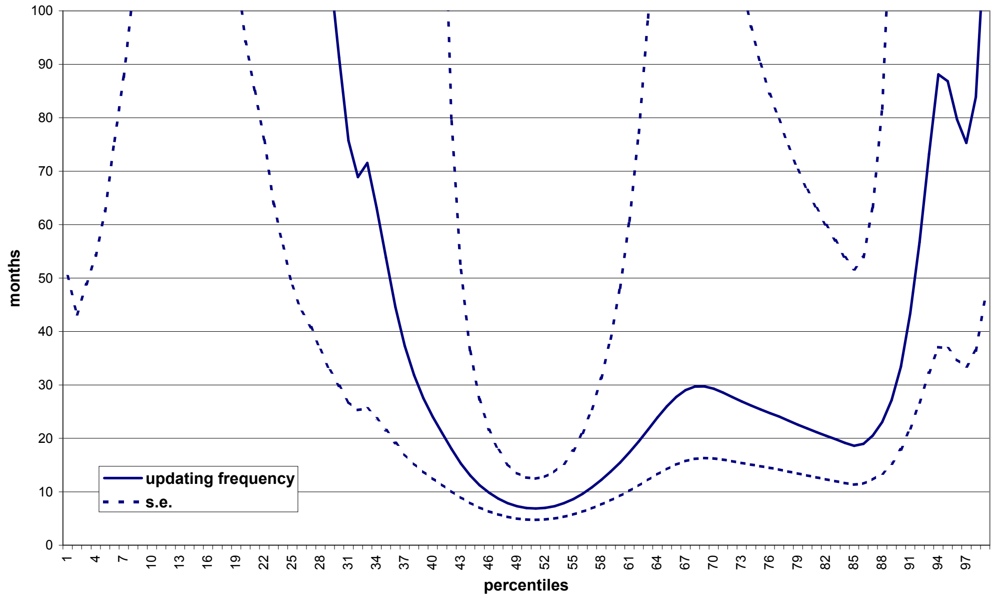


Figure 31:

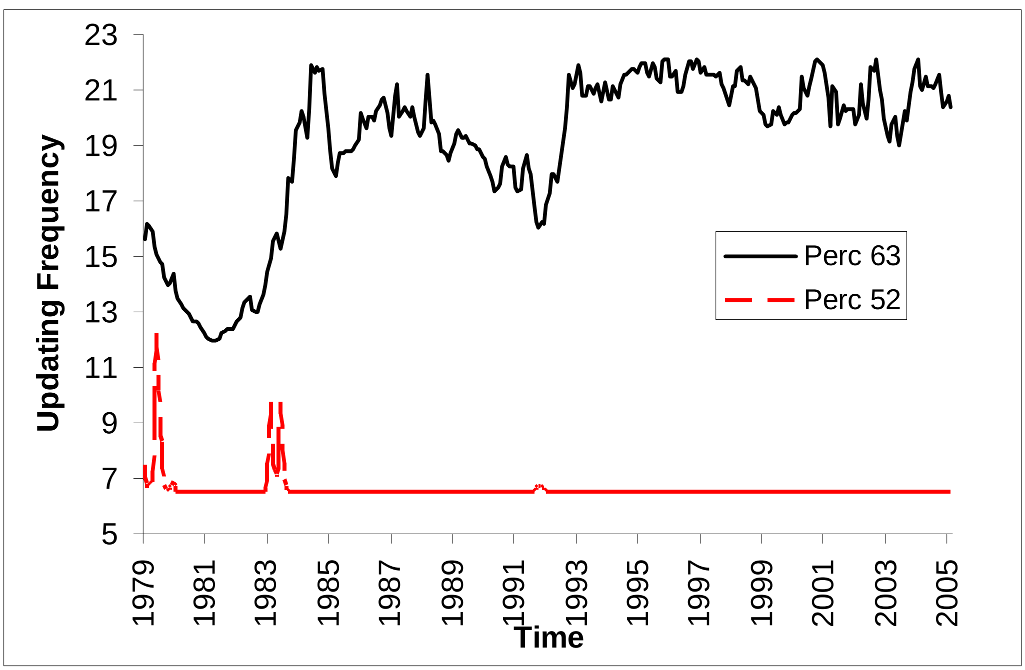


Figure 32:

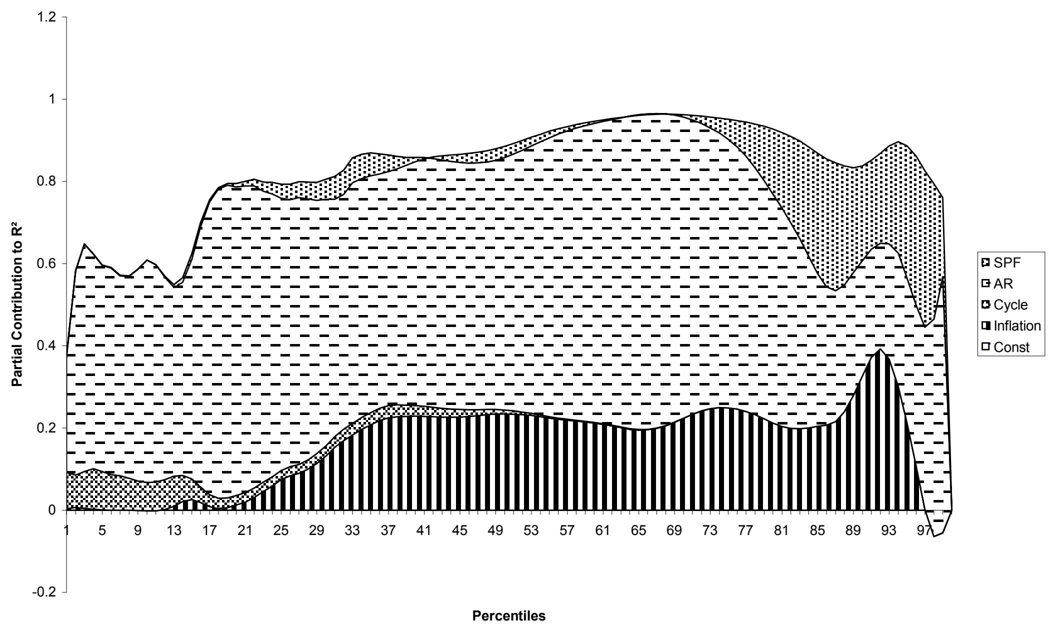


Figure 33:

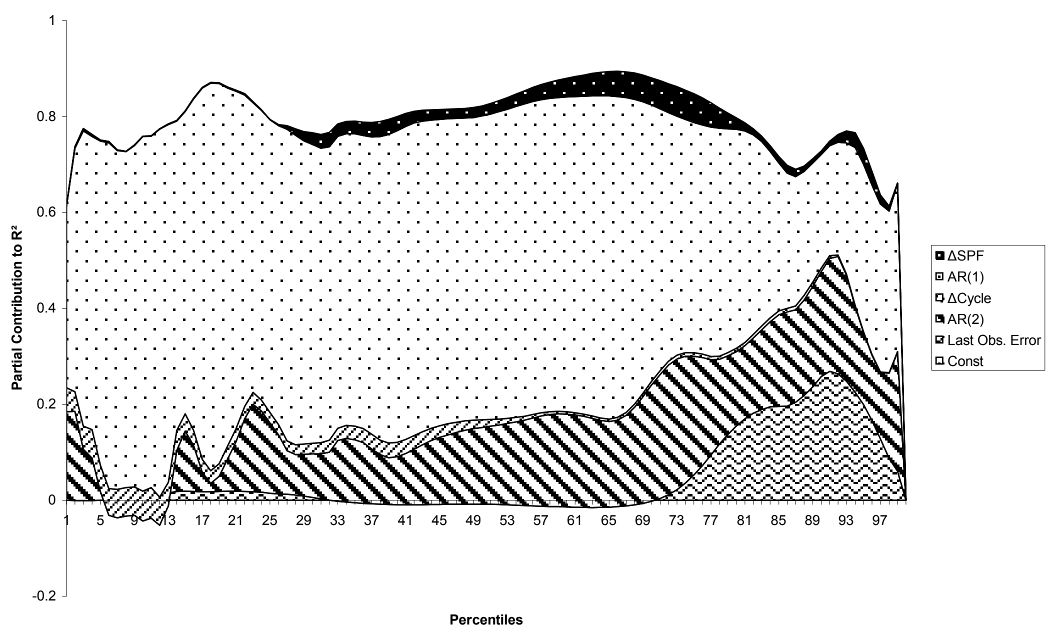


Figure 34:

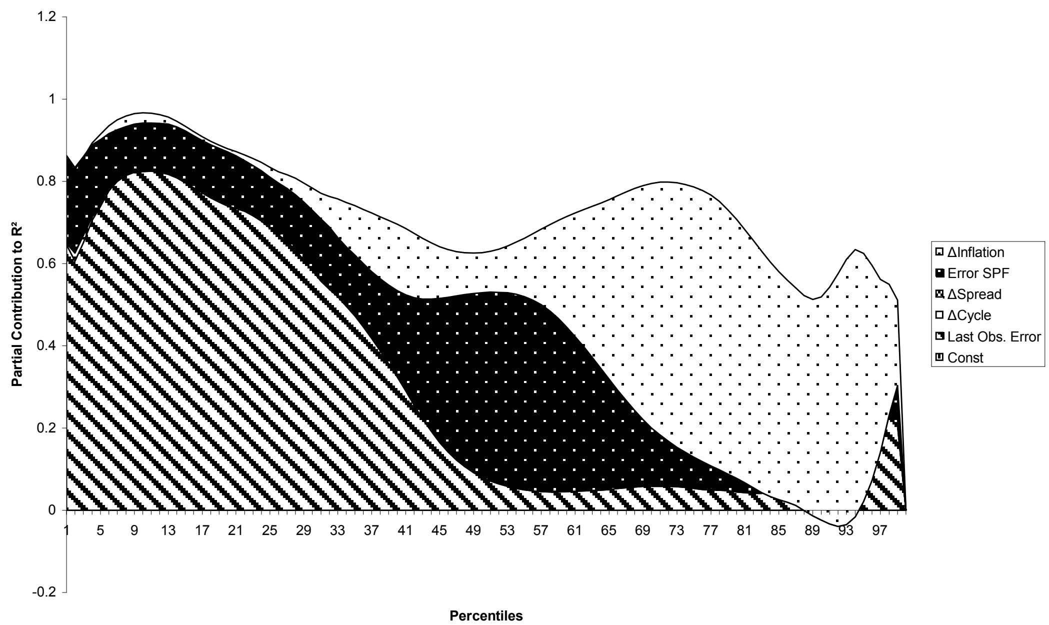


Figure 35:

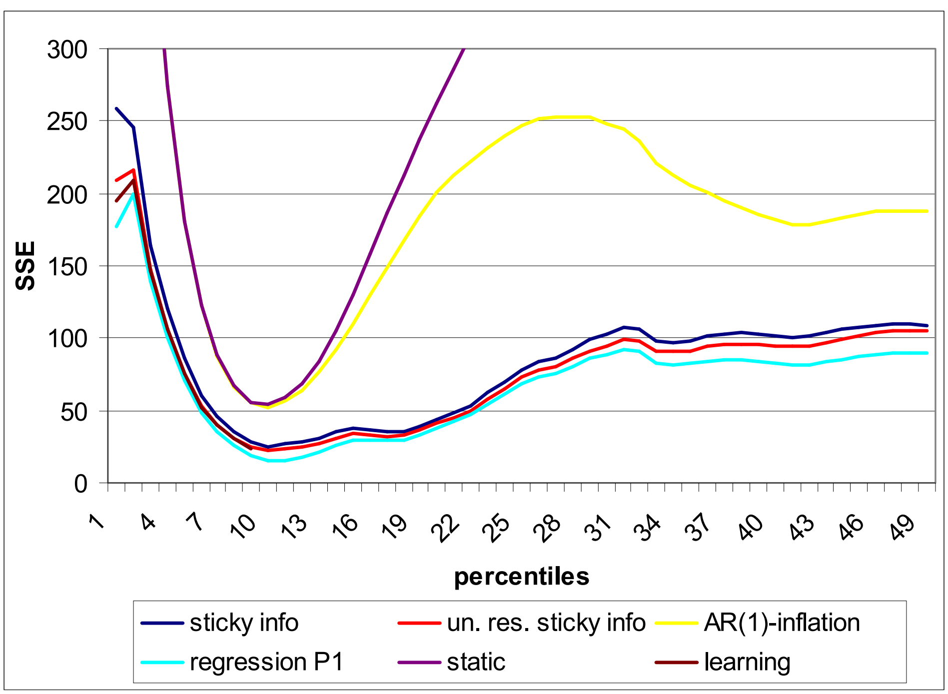


Figure 36:

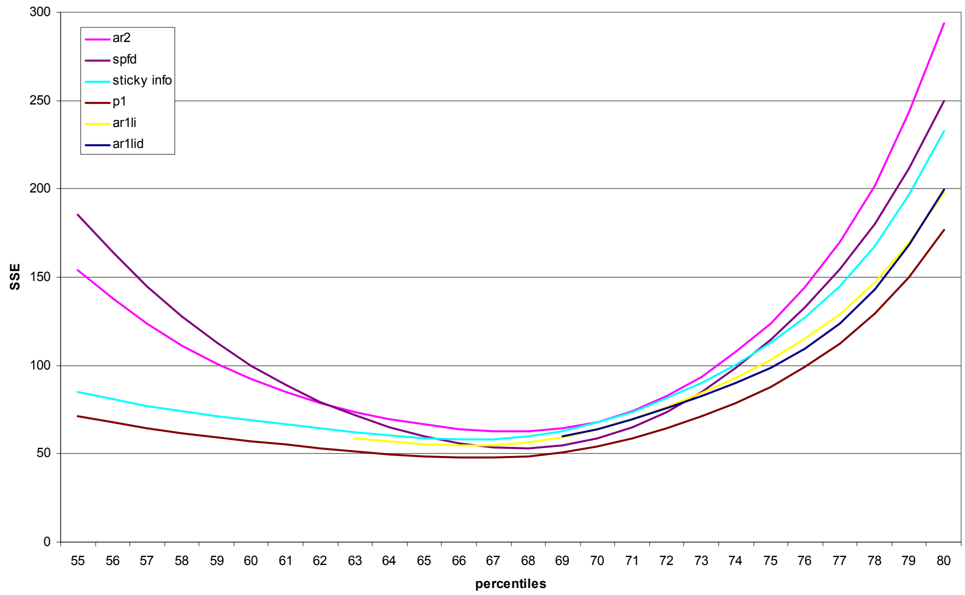
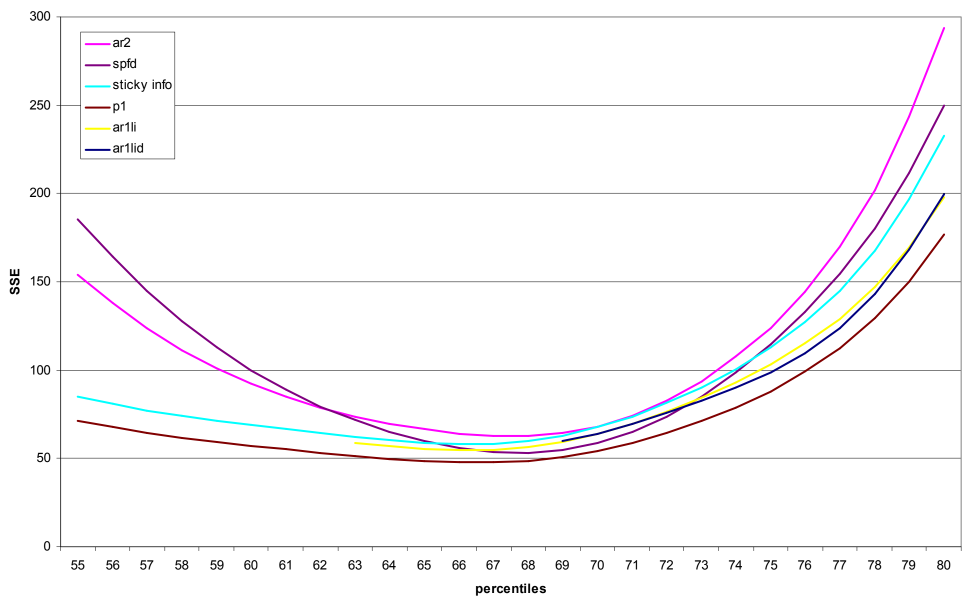


Figure 37:



6 Appendix B: Percentile time series analysis

[To be added]

7 Appendix C: Moments regressions

Table C1:

1st Model-pre1988	Constant	Trend	AR(1)	Cycle	Inflation	Term Str	R^2	DW
Interquantile Range	2.415	-0.010	0.604	-0.045	0.180	0.051	0.782	2.288
	2.994	-1.830	8.893	-0.487	2.716	0.464		
Variance	15.219	-0.035	0.415	-0.629	1.595	0.055	0.824	2.053
	4.447	-1.675	5.451	-1.558	4.732	0.120		
Skewness	0.908	0.004	0.345	0.007	-0.057	0.028	0.943	2.061
	6.367	5.942	4.358	0.727	-5.696	2.285		
Kurtosis	2.433	0.015	0.496	-0.015	-0.140	-0.032	0.895	2.079
	4.621	4.629	6.889	-0.313	-3.705	-0.555		

2nd Model-pre1988	AR(1)	Cycle	Inflation	Term Str	R^2	DW
Interquantile Range	0.371	0.004	0.234	-0.039	0.349	2.149
	4.389	0.054	3.536	-0.433		
Variance	0.223	-0.806	2.105	-0.874	0.559	2.010
	2.552	-2.488	6.131	-2.344		
Skewness	0.323	0.012	-0.058	0.043	0.631	2.093
	3.719	1.328	-5.336	3.672		
Kurtosis	0.448	-0.019	-0.132	0.041	0.458	2.093
	5.436	-0.472	-3.425	0.840		

Table C2:

1st Model-post1988	Constant	Trend	AR(1)	Cycle	Inflation	Term Str	R^2	DW
Interquantile Range	2.201	-0.001	-0.155	0.215	0.124	0.307	0.392	2.083
	4.604	-0.655	-3.289	3.454	2.585	4.376		
Variance	12.122	-0.026	-0.610	2.122	0.894	0.350	0.717	2.042
	4.650	-3.882	-2.445	5.765	3.449	5.143		
Skewness	2.061	-0.001	0.034	-0.109	-0.044	0.313	0.335	2.152
	8.094	-2.854	2.085	-4.607	-2.594	4.494		
Kurtosis	11.454	-0.002	0.248	-0.923	-0.422	0.109	0.445	1.946
	9.332	-1.002	2.833	-6.818	-4.413	1.489		

2nd Model-post1988	AR(1)	Cycle	Inflation	Term Str	R^2	DW
Interquantile Range	-0.121	0.223	0.106	0.313	0.224	2.065
	-2.668	2.823	1.549	4.493		
Variance	-0.700	2.893	0.808	0.278	0.416	2.006
	-3.114	6.663	2.421	4.117		
Skewness	0.025	-0.050	-0.015	0.243	0.115	2.068
	1.842	-2.071	-0.728	3.405		
Kurtosis	0.188	-0.608	-0.184	0.141	0.168	1.955
	2.489	-4.356	-1.573	1.952		