EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 721 / FEBRUARY 2007

ARE SURVEY-BASED INFLATION EXPECTATIONS IN THE EURO AREA INFORMATIVE?

by Ricardo Mestre

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 721 / FEBRUARY 2007

ARE SURVEY-BASED INFLATION EXPECTATIONS IN THE EURO AREA INFORMATIVE? [1]

by Ricardo Mestre [2]

In 2007 all ECB publications feature a motif taken from the €20 banknote.

1 This paper was presented in the CESIfo conference on Survey Data in Economics, 14th/15th September 2005. The author is grateful for comments from the discussant of the paper, Steffen Henzel, to M. Hashem Pesaran for an insightful discussion, and to the attendees to the conference for points raised in the general discussion. The opinions expressed are not necessarily those of the ECB.

2 Research Department, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany; e-mail: ricardo.mestre@ecb.int; Tel: +49 69 1344 7631; Fax: +49 69 1344 6575

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.int

Fax
+49 69 1344 6000

Telex
411 144 ecb d

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

This paper contributes to the old theme of testing for rationality of inflation expectations in surveys, using two very different surveys in parallel. Focusing on the euro area and using two well-known surveys that include questions on inflation expectations, the Consensus Forecast survey and the European Commission Household survey, a battery of tests is applied to inflation forecasts. Tests are based on a preliminary discussion of the meaning of Rational Expectations in the macroeconomic literature, and how this maps into specific econometric tests. Tests used are both standard ones already reported in the literature and less standard ones of potential interest within the framework discussed. Tests focus on in-sample properties of the forecasts, both in static and dynamic settings, and in out-of-sample tests to explore the performance of the forecasts in a simulated out-of-sample setting. As a general conclusion, both surveys are found to contain potentially useful information. Although the Consensus Forecasts survey is the best one in terms of quality of the forecasts, rationality in the European Commission Household survey, once measurement issues are taken into account, cannot be ruled out.

Key words: Rational Expectations, Tests of Rationality, Inflation Forecasting.

JEL:.C40, C42, C50, C53, E37

NON-TECHNICAL SUMMARY

The current paper analyses the degree of rationality of inflation forecasts as collected in two well-known surveys spanning the euro area. In doing so, two main objectives are sought: in the first place, putting some of the rationality tests commonly used in the literature in perspective in the light of euro area evidence; in the second place, report some results for surveys covering the euro area, in the hope of bringing the analysis on the issue closer to the level attained in the US. As a fall-out, it is expected that results reported in the paper will clarify the usefulness of these surveys.

The surveys used in the paper cover the euro area, as said, and are probably extreme cases in terms of the capabilities of the respondents: the Consensus Forecast (CF) survey and the Consumer survey by the European Commission (EC). The first is among the best in terms of the professional level and knowledge on the topic by respondents. The second one collects views of consumers that only casually think about inflation. The stark contrast between the two surveys strengthens the analysis. It may be expected that results would trivially point to strong performance of the professional forecasters and poor performance of consumers. One of the main goals of the paper is to assess the degree to which this fact is present in the data. Rationality is an elusive concept, as will be seen, but it may be easier to assess the relative degree of rationality of two sets of agents. This paper may thus be seen as testing the rationality of consumers vis-à-vis the best professional forecasters, rather than speaking about rationality in absolute terms.

The analysis is non-standard in that two surveys have been used: the European Commission Consumer survey (EC), directed to consumers, and the Consensus Forecasts survey (CF), directed to professional forecasters. The two surveys are without doubt at the extremes of what can be expected of agents in terms of forecasting performance. This fact will be used to highlight the information in the forecasts of one survey relative to the other, rather than in absolute terms.

The very different format of the two surveys has somehow limited the analysis, which has been done only in terms of comparable tests that could be run across surveys. In particular, the consumer survey needed quantifying of the forecasts (which are only qualitative), which implied some care in the analysis of this survey. The quantification has been done using relatively standard techniques (due to Berk (1999), an extension of the techniques of Carlson and Parkin), which introduce some uncertainty as to the actual level of expected inflation respondents had in mind when answering. In order to check this uncertainty, a looser approach needs to be taken to test for rational expectations, to account for mis-measurement. This is done in the paper by proposing looser definitions of rationality: by letting mild forms of unbiasedness and auto-correlation to be present in derived expectations. No such departures are assumed for the CF survey.

Furthermore, two different types of tests are run: the first one for the ability of respondents to forecast inflation; the second one to actually test for rationality. Without sufficiently good forecasts, it is impossible to check whether there was rationality in them. Surprisingly, both consumers responding to the EC survey and experts responding to the CF survey seem to make sufficiently good forecasts to enable testing for rationality. In terms of testing for rationality itself, the two surveys are different in the quality of the information they disclose, but both contain some useful information.

Rationality has been tested along three dimensions: i) the extent of absence of bias in the forecasts; ii) the degree to which agents could hope to forecast inflation using their models; and iii) the quality of the forecasts compared to other benchmark (parametric) models.

The conclusions reached are:

– Models used by survey respondents (as they were measured in a so-called forecastability test) were of a sufficient quality to enable them to actually build good forecasts, without a sharp distinction between the two surveys;

– Forecasts themselves seem to be more problematic, in particular the EC survey: bias seems to be present and there is little evidence that past forecasts errors are corrected.

– Survey forecasts do not fare badly compared to simple parametric alternative models used in the paper, the point applying equally well to the EC and CF surveys.

In summary, although the paper has not proved nor disproved that agents in the surveys are rational, at least it can be argued that consumers do not fare that bad compared to professional forecasters. Although the former provide much better forecasts than the latter, as far as the basic ingredients entering rationality are concerned, information in the EC survey should not be ignored.

1. Introduction

The current paper analyses the degree of rationality of inflation forecasts as collected in two well-known surveys spanning the euro area. In doing so, two main objectives are sought: in the first place, putting some of the rationality tests commonly used in the literature in perspective in the light of euro area evidence; in the second place, report some results for surveys covering the euro area, in the hope of bringing the analysis on the issue closer to the level attained in the US. As a fall-out, it is expected that results reported in the paper will clarify the usefulness of these surveys.

The surveys used in the paper cover the euro area, as said, and are probably extreme cases in terms of the capabilities of the respondents: the Consensus Forecast (CF) survey and the Consumer survey by the European Commission (EC). The first is among the best in terms of the professional level and knowledge on the topic by respondents. The second one collects views of consumers that only casually think about inflation. The stark contrast between the two surveys strengthens the analysis. It may be expected that results would trivially point to strong performance of the professional forecasters and poor performance of consumers. One of the main goals of the paper is to assess the degree to which this fact is present in the data. Rationality is an elusive concept, as will be seen, but it may be easier to assess the relative degree of rationality of two sets of agents. This paper may thus be seen as testing the rationality of consumers vis-à-vis the best professional forecasters, rather than speaking about rationality in absolute terms.[1]

The topic has been of constant interest for the ECB, see e.g. the boxes included in its Monthly Bulleting issues of July 2002, October 2003 or April 2005. Furthermore, a full article was dedicated to the topic of measures of inflation expectations in the euro area, see the July 2006 issue. At a more technical level, Forsells and Kenny (2002) include an assessment similar to this paper, although the looser approach taken in the current text partially reverses their conclusions.

The next section will discuss methodological issues necessary for a better understanding of the problem, including what should be understood by rational expectations in this context and how to test for them. Rather than discussing from a broad perspective, focus will be on the mapping of concepts and the empirical information in the surveys. Section 3 will then briefly present the data. Section 4 will tackle an analysis of inflation expectations in the surveys in terms of testing the rationality and forecasting ability of respondents, both using in-sample and simulated out-of-sample tests. Section 5 will conclude.

An appendix describes the surveys and some technical aspects of the approach followed.

[1] The format of the two surveys is unfortunately very different, and this has limited somewhat the analysis. There is clearly a trade-off between comparing the two surveys and working with them separately.

2. The Framework and the Methodology

It is necessary to start any discussion on rationality by first defining the concept and looking into ways to test it. The approach taken in the analysis is the fairly standard one of linking choices of economic agents to their understanding of the economy (i.e., their awareness of the model underlying the economy) and their expectations about future events (i.e., their projections using the model). As is standard in the literature, and probably is unavoidable due to stringent information requirements otherwise, it will be assumed that economic agents use reduced-form, simple models—maybe informal ones—to draw their own plans. Rational expectations in this setting is not necessarily about knowing the economy, it may be about getting things 'right' on average but not always.

2.1 Theoretical Motivation

Let's assume that a log-linearised macro model is as expressed in (1), where $x_t - \bar{x}_t$ is a representation of deviations from steady-state values and ε_t an unforecastable residual assumed momentarily to be non-autocorrelated. Correspondingly, the steady state is described by a number of relationships in (2), which will be assumed unrelated to business-cycle behaviour in (1). In (1), not all variables in x_t will be observed, but the agents are assumed to know which observables are relevant.

$$A_{-1}\,\mathrm{E}_t\left(x-\bar{x}\right)_{t+1} + A_0\left(x-\bar{x}\right)_t + A_1\left(x-\bar{x}\right)_{t-1} = B_0\,\varepsilon_t \tag{1}$$

$$f\left(\bar{x}_t, \bar{x}_{t-1}\right) = 0 \tag{2}$$

Agents are assumed to form expectations according to the reduced-form model (3) for the business-cycle component and (4) for the steady state.

$$\mathrm{E}_t\left(x-\tilde{x}\right)_{t+1} = C_1\left(x-\tilde{x}\right)_t \tag{3}$$

$$g\left(\tilde{x}_t, \tilde{x}_{t-1}\right) = 0 \tag{4}$$

Assuming for one moment that agents understand correctly steady-state relationships (i.e., $\bar{x}_t = \tilde{x}_t$) and that there are no structural breaks, substituting out (3) in (1) and solving yields (5). If agents are rational, the forecasting model (3) and the reduced-form model (5) should be observationally equivalent. Under the assumption that (1) admits one single solution, agents will in this case eventually learn that C_1 is such that it solves $C_1 = -\left(A_{-1}C_1 + A_0\right)^{-1}A_1$ and is compatible with a stable solution of system (1).[2]

$$\left(x-\bar{x}\right)_t = -\left(A_{-1}C_0 + A_0\right)^{-1}A_1\left(x-\bar{x}\right)_{t-1} + \left(A_{-1}C_0 + A_0\right)^{-1}B_0\,\varepsilon_t \tag{5}$$

This leads to a number of interesting properties a rational-expectations solution should have:

1. Expectations should differ from actual values by an unforecastable residual.
2. Expectations should be formed using all relevant information in the available data set, i.e. all observed pre-determined variables that matter for the model solution should enter the expectations-formation mechanism and nothing more.

[2] In case of multiple equilibria, we will assume that agents are able to learn the actual equilibrium reached.

3. Expectations should be efficient, in the sense that alternative forecasts should lead to errors with higher variance than rational expectations.

It should be evident that the amount of information and processing capability required from the agent is staggering. This extreme definition of the rational expectations hypothesis (REH), in which agents know, even approximately, all relevant aspects of the economy and are able to exploit them, is what we will term *strong* version of REH.

The econometrics of the problem turns to be somewhat more interesting if, unbeknown to the agent, the residual in (1) is autocorrelated, as seen in (6). In this case, the implicit reduced-form residual in (5) would be autocorrelated itself, and the forecaster would have to consider some dynamics in its reduced-form regressions in order to estimate efficiently.

$$A_{-1}\,\mathrm{E}_t\left(x-\overline{x}\right)_{t+1} + A_0\left(x-\overline{x}\right)_t + A_1\left(x-\overline{x}\right)_{t-1} = B_0\,\varepsilon_t - B_1\,\varepsilon_{t-1} \tag{6}$$

In case there is autocorrelation not fully detected by the forecaster, condition 1 above should be replaced by:

1'. The difference between expectations and actual values should be a stationary zero-mean process.

Furthermore, condition 3 should be qualified, since off-model information (in the form of lags of endogenous variables entering the forecasting model in seemingly non-structural ways) is needed.

Things are a bit more complex if the assumption of perfect knowledge of the steady-state conditions is dropped. In this case, errors can be made in the long-run conditions, which in turn will affect the business-cycle perceived conditions. Assuming that steady-state conditions are not observed but derived through filtering observed variables, as in $\tilde{x}_t = G\left(L\right)x_t$, then any error in the filter will induce auto-correlation in the residuals in (5) and maybe also *non-persistent* biases in the forecasts.[3] The same conclusion applies if structural breaks happen that need to be learned. In particular, if the steady-state conditions are set for non-stationary variables, under rational expectations actual and forecasts should be co-integrated.

This is what we will call in the sequel *weak* form of REH. Note that a weak version of REH simply amounts to allowing for some degree of error auto-correlation in the forecasts and maybe also some amount of bias in forecasting. We will still call this situation REH because we are maintaining the assumption that agents (i.e., forecasters) understand the basic elements of the model underlying the economy.[4] This understanding will be assumed to rest on successful fitting of the data, rather than a full understanding of the structural intricacies in the economy.

The goal of this paper will be to analyse empirical counterparts of (3), based on explicit or implicit expectations recorded in the surveys.

---

[3] A rational agent should eventually learn how to correct for past biases.

[4] Note that parameters in (3) could be time-varying, which induces yet another layer of complexity to the problem. This point, though, will not be taken up in the text.

2.2 The Econometric Representation

Let's say $\pi_t^h(i)$ is a representation of agent's i expectation of inflation for period t, h periods ahead, taken from the survey. The corresponding survey average will be termed π_t^h, where again the superscript h stands for the forecast horizon, i.e. forecasts collected h periods before t. Note that both h and t will be monthly for the EC survey, but that t will be annual and h monthly for the CF, as forecasts in this survey are for full years.

The basic assumption made in the text is that respondents in both surveys report the mean forecast coming from their own 'view' of the world (i.e., model) and the information available to them. In other words, respondent i uses a model for inflation to build expectations: $\pi_t^h(i) = \mathrm{E}\left[\pi_t \mid \Omega_{t-h}(i), M_{t-h}^h(i)\right]$, where $\Omega_{t-h}(i)$ is the information set used and $M_{t-h}^h(i)$ is the model used to produce the forecast, both potentially different for each forecaster. The respondents don't report any other measure of the distribution of their forecasts. This implies that forecasts will differ among respondents because of differences in information sets and/or models, but this variability in the forecasts will bear no relationship to the perceived level of actual uncertainty in the economy.

Note also that there is some ambiguity about what kind of inflation is being forecasted: it could be a broad definition, observable by others, or a local concept affecting each forecaster in particular. The CF survey is clear in asking respondents about country inflation rate for consumer prices (for the euro area inflation as of 1999). The EC survey is a lot less clear about what is being asked (see the annex for the drafting of the questions). It could be envisaged that consumers are actually answering about their perceptions of inflation for the bundle of goods they usually purchase, i.e. an extremely local concept. This possibility has affected the analysis below.

Obviously, the actual outcome for inflation will depend on fundamentals driven by the 'true' model of the economy plus an unforecastable shock, equation (7), where $\mathrm{E}_{t-h}\pi_t$ is the forecast that would have obtained had the forecasters known the true model, and υ_t^h is a weakly stationary, zero-mean process uncorrelated with $\mathrm{E}_{t-h}\pi_t$ and auto-correlated at most h-1 periods. Note that it is infeasible to use the 'true' model because of information-processing costs and other difficulties in doing so. The true model may seem an uninteresting theoretical construct, but it in fact affects rationality tests since it sets an upper bound to the quality of the forecasts: as will be shown, it is futile to test for rationality if inflation is basically unforecastable.

$$\pi_t = \mathrm{E}_{t-h}\pi_t + \upsilon_t^h \qquad (7)$$

It will be assumed that rational forecasters use proxies to the true model that bring an optimal mix of forecasting power and ease of use. Not all the forecasters will choose exactly the same proxy model, though, so there should be some variation in the forecasts even in the case of rational forecasters. If REH holds, forecasters will be on average close to the true model. We can describe the situation by assuming that forecasts are formed using (8), where the error process μ_t^h, different from υ_t^h in the true model, will be assumed weakly stationay, zero-mean and auto-correlated of at most order h-1. This error differs from υ_t^h to the extent that agents use an approximation to the true model, and so includes model (and parameter) uncertainty. It could obviously happen that agents roughly coincide on the model they want to use, but the chosen model is very different from the true one, in which case there would be no rationality. But, it may be argued, it seems unlikely that professional forecasters keep on using wrong models, which should give inferior performance. Consumers may be more prone to this kind of mistake, but we are at least in a position to assess their behaviour relative to the performance of professional forecasters.

$$\pi_t^h = \mathrm{E}_{t-h}\pi_t + \mu_t^h \tag{8}$$

With these definitions at hand, it is trivial to see that REH entails both that forecasters use a 'sufficiently' good approximation to the actual data-generation process (DGP) of inflation, and that the DGP is 'sufficiently' good to actually lead to informative forecasts, in a sense that will be made clearer below. In other words, forecasters spend time and effort in getting a good forecasting model because there are gains to be had from using this model that outweigh the costs of building and operating it.

$$\pi_t = \alpha + \beta\,\pi_t^h + \varepsilon_t^h \tag{9}$$

Note that if (7) and (8) both hold, a testing regression such as (9) can be used to test for rationality, as is actually the most common case in the literature. One of the earliest and most popular tests in the literature is the common test that $\alpha = 0$ and $\beta = 1$ in the equation. It can be easily shown that, under both (7) and (8), the estimate for β is given by:

$$\hat{\beta} = \frac{\mathrm{var}\left(\mathrm{E}_{t-h}\pi_t\right) + \mathrm{cov}\left(\upsilon_t^h, \mu_t^h\right)}{\mathrm{var}\left(\mathrm{E}_{t-h}\pi_t\right) + \mathrm{var}\left(\mu_t^h\right)},$$

which will fail the test if υ_t^h and μ_t^h are very different—for instance, if both are weakly uncorrelated or if the latter is much larger than the former. Since error μ_t^h is related to the errors made in approximating the 'true' model for inflation, this test is rightfully testing the degree of rationality of the forecasts.

$$\pi_t^h = \alpha + \beta\,\pi_t + \varepsilon_t^h \tag{10}$$

Another matter is whether the 'true' model of inflation is sufficiently good to actually serve any useful purpose. It may be that the forecast error υ_t^h is so large that forecasts are in practice useless. This can be tested, using the forecasts under analysis, by running the regression in (10) and testing again whether $\alpha = 0$ and $\beta = 1$. In this case,

$$\hat{\beta} = \frac{\mathrm{var}\left(\mathrm{E}_{t-h}\pi_t\right) + \mathrm{cov}\left(\upsilon_t^h, \mu_t^h\right)}{\mathrm{var}\left(\mathrm{E}_{t-h}\pi_t\right) + \mathrm{var}\left(\upsilon_t^h\right)},$$

and the test will fail if both are uncorrelated or if υ_t^h is large enough relative to μ_t^h. This test could be described as testing the forecastability of inflation.[5] Obviously, if inflation is not forecastable with the models used by the forecasters, it is meaningless to test for rationality. The problem with a forecasting

[5] Note that we use the term forecastability in a slightly loose way.

test, which will be only briefly touched upon in this paper, is that forecastability is a relative concept: it should be assessed only with respect to the best possible forecasting model. In the text below, some alternative benchmarks will be mentioned, but this point will not be explored in any depth.

2.3 A caveat

It may be worth reminding the already old issue about whether to average responses or to pool them, see e.g. Dietrich and Joines (1983). In this paper, both surveys will be used following a similar approach for the two: only averages will be considered. Furthermore, forecasts will be judged against euro area inflation, i.e. without distinguishing between different inflation rates in the member countries.

In what respects the EC survey, there is little choice but to average forecasts due to the format of the answers: as they are of a qualitative nature, averaging is necessary to extract a quantitative measure. For the CF survey, averaging responses is done basically because of the ultimate goal of the paper of comparing the two surveys. There is a growing literature on the improvement of forecasting performance of combining different forecasts, see e.g. Hendry and Clements (2001). Since we are not interested in testing rationality on an individual basis, but rather whether forecasters have a sophisticated view of the economy *on average*, we proceed to average responses.

Averaging is also needed because of specificities of the data used in the analysis. Take the EC survey, for instance. As already indicated, the drafting of the questions on inflation expectations does not rule out that respondents are being asked about their local inflation rates. If so, the only meaningful analysis is by first aggregating and averaging survey responses. The rationale in this case is that euro area inflation as collected by statisticians will be calculated as a weighted average of local inflation rates, and that consumer idiosyncratic inflation rates and local inflation rates as collected by statisticians will be similar, and weights in data collection and survey sampling will also be similar. These two assumptions should hold if both the surveys and data collected by statisticians are representative of the population. One consequence of this kind of forecasts is that it may well happen that data observed by the analyst and data used by the forecaster differ. This in turn implies that the efficiency of use of the original information set cannot be tested, since the analyst making the test does not observe it fully.

A similar point applies to the CF survey, which for most of the sample was collected for country inflation rate expectations. Although less extreme a case than for the EC survey, caution is needed if analysing individual answers in a study done at the euro area level.

Last but not least, averaging is also done to avoid biases in the tests introduced by poor forecasters. Note, for instance, that if tests are based on the regression $\pi_t = \alpha + \beta \pi_t^h(i) + \varepsilon_t^h$, as is often done in the literature, it is trivial to show that under REH $\mathrm{E}\alpha = 0$ and $\mathrm{E}\beta = \mathrm{var}\left(\pi_t^h(i)\right) / \left(\mathrm{var}\left(\pi_t^h(i)\right) + \mathrm{var}\left(\nu_t^h\right)\right)$, which implies a downward bias for the slope coefficient. Averaging across respondents smooths out the resulting forecast and reduces the bias.

2.4 Rational expectations

It is necessary to agree on a definition of what rational expectations are from an econometric-testing perspective. As seen previously, two basic ingredients are needed from the theoretical point of view: unbiasedness and efficiency. The first basically entails that inflation and its expectations should co-move, the second that no extraneous information should lead to better forecasts of inflation. Loosely speaking:

$$\mathrm{E}\left(\pi_t - \pi_t^h\right) = 0\,,$$

$$\mathrm{var}\left(\pi_t - \mathrm{E}\left(\pi_t \mid \Omega_{t-h}\right)\right) \geq \mathrm{var}\left(\pi_t - \pi_t^h\right),$$

where π_t^h is as before a representation of agents' rational expectation of inflation *h* periods ahead and $\mathrm{E}\left(\pi_t \mid \Omega_{t-h}\right)$ is an alternative forecast based on a given model and information set proposed by the analyst. Note that both conditions should optimally hold, but that none must do it strictly in the case of weak rationality: the first condition could be violated momentarily in the case of not-corrected biases; the second could be violated in case of remaining dynamics in the forecast errors.

In terms of testing mechanisms, we want to describe and test the statistical properties of the forecast errors implicit in the rational expectations. For this, it is necessary to distinguish between a strict (or strong) and a loose (or weak) version of the concept, in the sense defined above. Additionally, it is important to adapt the analysis to the number of unit roots present in the system, i.e. whether the two variables are stationary, whether there is one unit root (co-integration) or whether there are two unit roots.

We will define strong rational expectations as in (11) in case series are stationary or co-integrated, and (12) otherwise.

$$\pi_t - \pi_t^h = \varepsilon_t, \quad \mathrm{E}\left(\varepsilon_t \mid \Omega_{t\text{-}h}\right) = 0 \tag{11}$$

$$\Delta\pi_t - \Delta\pi_t^h = \varepsilon_t, \quad \mathrm{E}\left(\varepsilon_t \mid \Omega_{t-h}\right) = 0 \tag{12}$$

Similarly, weak rationality will be defined by (13) and (14), in which the error process is no longer white noise, although it remains stationary.

$$\pi_t - \pi_t^h = \theta\left(L\right)\varepsilon_t, \quad \mathrm{E}\left(\varepsilon_t \mid \Omega_{t\text{-}h}\right) = 0, \quad \left|\theta\left(z\right)\right| < 1 \tag{13}$$

$$\Delta\pi_t - \Delta\pi_t^h = \theta\left(L\right)\varepsilon_t, \quad \mathrm{E}\left(\varepsilon_t \mid \Omega_{t\text{-}h}\right) = 0, \quad \left|\theta\left(z\right)\right| < 1 \tag{14}$$

Note that none of the tests performed below use differenced inflation, but a distinction between static and dynamic tests will be made. Dynamic tests are robust to lack of co-integration, in case series are I(1).

Note that one important implication of weak rationality is the fact that lagged information may matter in a test, contrary to the strong case.

Last but not least, although it can be accepted to have some small bias, a large and persistent bias is not acceptable.

2.5 Different types of tests

The tests proposed and performed in this paper will follow the definitions given above:

i) Firstly, tests of REH in its strong and weak definition.

ii) Also, test for forecastability.

iii) Test the forecasting ability of the survey respondents.

Note that an analysis of efficiency is made more difficult by the nature of the surveyed expectations. Both the EC and CF surveys address expectations of inflation only indirectly for the euro area. It is thus difficult to gauge what data forecasters were using when making their forecasts. It can be assumed that CF forecasters were using country information, at least before 1999, but EC respondents were probably using data simply not observed by the analyst. As a consequence, no attempt has been made to analyse orthogonality conditions between forecast errors and data available at the time. All the tests are, implicitly or explicitly, whether survey forecasters exploited all information contained in past inflation.

3. Data

Data for inflation relate to euro area HICP. Since series for HICP published by Eurostat only begin in 1991, some degree of backdating is necessary. This has been achieved by linking the aggregation of national CPIs for the previous periods to the series published by Eurostat. Data for inflation expectations are from the EC and CF surveys, using methods described in the Appendix. The sample differs for each survey: 1985 to 2003 for the EC survey, 1990 to 2003 for the CF survey.

CF forecasts were taken in raw form. They correspond to monthly forecasts of current-year inflation and following-year inflation.

EC Consumer forecasts are reported in qualitative form, and must somehow be quantified. The questions in the survey refer to current perceived inflation and to expected inflation relative to current one (see annex). The quantification of the forecasts is made in two ways: the first one, by assuming that consumers refer their statements on expected inflation to last month's actual inflation, i.e. assuming that they are aware of past inflation and base their forecast on it; the second one, by assuming they refer to a subjective 'perceived' inflation which must be quantified from another question in the survey.

An annex to the paper thoroughly describes both the questions in the survey, the procedures followed to quantify the EC survey and also chart the forecasts.

4. Expectations as indicators of inflation

As mentioned above, a successful set of inflation expectations should send useful signals about future inflation. In other words, these forecasts should provide unbiased and low-variance estimates of yet-to-be-released inflation data. There are two broad sets of tests that can be used to test this: in-sample tests, using the full sample to derive estimates of the links between expectations and inflation that could lead to answers; and out-of-sample tests, in which the analyst tries to replicate a situation in which forecasts are made exploiting no information coming from the forecasted period or later.

4.1 In-sample properties

Usually, the initial question when dealing with data is their stationarity properties. Series used in this paper for inflation and inflation expectations are seemingly I(1) on statistical grounds.[6] Theory would rather support them being I(0), at least in the latter part of the sample. The approach taken in this paper has been in this respect a pragmatic one: the paper includes a battery of tests standard in the literature, some of which are sensitive to the stationarity assumption. Tests affected by lack of robustness to non-stationary behaviour have been properly identified in the text, and alternatives offered in each case that are robust to the assumption. But no stand is taken on whether inflation and inflation expectations are stationary or not.

4.1.1 EC Consumer survey forecasts

Question Q6 of the EC survey was first tested using the static- and dynamic-regression approach. Regressions with inflation on the LHS and a constant and expectations on the RHS were first run, i.e. the most standard test in the literature. The regressions were then reversed, i.e. expectations were put on the LHS and inflation on the RHS. These are tests using static equations. The tests were then repeated using dynamic equations. The dynamic equations were built in the following way:

i) Inflation was put on the LHS. On the RHS, initially, a constant, seasonal dummies, contemporaneous expectations, 12 lags of expectations and 12 lags of inflation were added. Since the number of regressors was high (1 constant, 11 seasonal dummies, 12 lags of the endogenous variable, 13 lags of the indicator), the equation was streamlined using PcGets. [7] The process was then repeated putting the forecasts on the LHS.

ii) The final regression was solved for the implicit long run of the equation, i.e. the sums of the parameters for the lags of the variables were added up.

In a nutshell, the estimated dynamic equation was:

$$\pi_t = \alpha + \sum_{s=0}^{12} \beta_i \pi_{t-s}^h + \sum_{s=1}^{12} \gamma_i \pi_{t-s} + seasonal + residual. \tag{15}$$

The test was for $\sum_{s=0}^{12} \beta_i + \sum_{s=1}^{12} \gamma_i = 1$ and $\alpha = 0$ jointly. Note that, due to PcGets, not all lags were actually present. Note also that the test is useful to check for co-integration if the two variables are I(1)[8]

[6] This statement is based on a set of ADF tests run on the series, using a variety of lag settings and deterministic components. As is well-known, ADF tests are particularly robust. Results are available on request.

[7] PcGets© is a module attached to GiveWin© and PcGive© that streamlines single-equation regressions, dropping unnecessary regressors while looking for stable and well-behaved equations. Note that the equation can be used without any streamlining; PcGets was used to better summarise the information contained in the regressions; no essential information was lost or misrepresented because of that.

[8] Let's stress, though, that it is not a formal co-integration test.

and is thus a useful complement to the static equations: if variables are I(1) and there is lack of co-integration, the static-regression parameters may be affected by spuriousness.

The same two tests were run for the two quantifications of the survey used in this paper: one building expectations from last month's inflation and the other building expectations on top of quantified perceived inflation.

Results are included in Table 1 (for expectations built from actual past inflation, variable ECQ6 in the table) and Table 3 (for expectations built from perceived inflation) for the rationality tests; and Table 2 and Table 4 for the corresponding forecastability tests. The variable measuring inflation in the tables was year-on-year HICP inflation rate (variable INFLATION in the box). Table 1 is a bit longer to better clarify the approach taken. In the table, the full regression results are first documented—reported is a rolling window regression and a number of standard tests. The second part of Table 1 presents the sum of dynamic parameters, the static solution to the equation (which is similar to a co-integration test) and the significance of each variable. In order to shorten the presentation, the rest of the tables do not report full regression results and focus instead on the sum of parameters.

Two points are worth noting.

1. In the first place, the relationship between the first measure (based on past observed inflation) and inflation is far from satisfactory, see tables 1 and 2. Although the sum of parameters of the LHS variable (INFLATION) is significantly different from zero, the same cannot be said for the sum of parameters of the forecasts. Furthermore, the static solution of the equation leads to a long-term relationship between the two variables of 0.50, far from 1. On the other hand, the forecastability test (with forecasts in the LHS) was much better, with clear indications of strong long-term relationship (a t-stat of around 10 for both the LHS and RHS variables) and a lack of bias in the slope parameter (the static solution leads to a long-term parameter of 1.05, not significantly different from 1). The constants in both tests are different from zero, so some bias is present.

2. The relationship between the second measure (based on perceived inflation) and inflation was much worse, see tables 3 and 4. The REH test is failed both in terms of bias and efficiency. The forecastability test, with forecasts in the LHS of the equation, was better but again worse than in Table 2. PcGets dropped the constant from the equation, so no bias was present.[9]

In summary, results in terms of rationality are poor for the first indicator, forecasts based on past inflation and very poor for the second. Forecastability tests are better, with the implication that respondent's to the survey ought to exploit their information more efficiently than they do.

[9] The constant was non-significant anyway, so this statement does not depend on the handling done by PcGets.

4.1.2 Consensus Forecasts

As for the CF Forecasts, only annual regressions can be run using the raw data, since forecasts themselves are for full years. In particular, surveys collected in the same month must be treated separately, since each entails a specific number of periods ahead for the expectations: in January, 12 and 24 months ahead; in February, 11 and 23 months ahead; etc. Unfortunately, this leads to a short number of observations: 13 for the current-year forecasts and 12 for the next-year forecasts. This implies that each test involves 24 separate equations, which makes the test impossible to report in table format. Instead, it will be reported in charts.

As before, tests involve parameters of both static and dynamic regressions. For each month by separate, regressions were run and parameters reported in graphical form. For instance, a regression was done for surveys in December and the constant, slope and joint significance test were graphed for 1-step-ahead forecasts (point 1 on the horizontal axis). The same was done for surveys in November, reported as 2-step-ahead forecasts, and so on. Next year forecasts were correspondingly reported as being 13- to 24-step-ahead forecast on the horizontal axis. The sequence of constants and slopes are reported with standard confidence bands, the joint significance test as its significance, such that significance levels above (say) 10% indicate successful tests. Results for REH tests are reported in Chart 1. REH tests based on dynamic equations were reported in similar way, except that the slope is reported as a double line: one line corresponds to the sum of parameters for the LHS variable, the other for the RHS variable. If the regression is expressed as $\theta(L)\pi_t = \alpha + \beta(L)\pi_t^h + \ldots$, lines depict the quantities $\theta(1)$ and $\beta(1)$. In this case, the test is not whether the slope is 1 but whether the two sums are identical. Dynamic results are shown on Chart 2.

Last but not least, Charts 3 and 4 show similar results for the forecastability tests, i.e. using forecasts as LHS variable.

Results point to somewhat better behaviour of the CF forecasts than was the case for the EC survey. Static REH tests are passed for the surveys run towards the end of the year, up until releases in August: the slope is not far from 1, the constant is not far from 0, and the joint test is passed, see Chart 1. Results for the next-year forecasts (panels on the right) are more mixed. Chart 2 reports similar tests for the dynamic equations. Again, there is evidence of rationality for the surveys towards the end of the year (this time, up until surveys in June), to be qualified because parameters sum is not significantly different from zero—observe the bands. The corresponding forecastability tests, Charts 3 and 4, are better behaved than the REH tests, just as was the case for the EC survey.

No attempt was made to look into alternative forecasting models, due to the already mentioned difficulties in assessing what data were used by the forecasters to build their forecasts.

Clearly, the evidence of rationality in the CF survey is much stronger than in the EC survey, which is hardly surprising. The CF forecasts seem to be a relatively good indicator tool, while the EC survey needs some further testing. The only caveat with the CF survey is problems found with the dynamics of the next-year forecasts.

4.2 Out-of-sample properties

This section tackles how well inflation expectations do predict inflation. While the focus in the previous section was on *in-sample* fit, the focus now will be on out-of-sample behaviour. The latter approach needs alternative indicators as benchmarks, since forecasting performance can only meaningfully be tested against alternative forecasting models. Although the topic of searching for optimal forecasting models is interesting on its own, the approach taken in this paper is a simpler one by which benchmark models will lack any indicator, only lags of inflation itself been included in the equations. The framework chosen will comprise two different forecasting tests: firstly, tests using the derived expectations series as raw inflation indicators; secondly, by using them as indicators in forecasting regressions. As will be explained below, using raw surveys answers is the appropriate way to test the forecasts in the surveys *provided that they are a correct measure of respondent's expectations*. As has already been said for the EC survey forecasts, numbers reported have been derived and may not represent consumers' views. For the EC survey, hence, it seems necessary to use the forecasts indirectly, in the form of indicators embedded in a forecasting equation. The CF survey is less apt to be subjected to this approach.

Note that forecasting models, when used, have been simple equations of inflation—actually, log of the first difference of HICP—on own lags, a constant and contemporaneous and lagged values of selected indicators whenever appropriate, as in (16). Lags in the equations have been selected ex-post, i.e. after testing different configurations, as opposed to using statistically-based methods such as the well-known BIC criterion. This approach was taken because of simplicity. The equations have additional seasonal dummies to take into account the lack of seasonal adjustment of the series, although they could be dropped without seriously affecting results.

$$\Delta\pi_{t+1} = A(L)\cdot\Delta\pi_t + B(L)\cdot z_t + \varepsilon_t \tag{16}$$

Tables 5 and 6 report results for the EC forecasts based on previous-period inflation, as the forecasts based on perceived inflation were clearly worse. Table 5 reports results using the survey forecasts in raw form. Table 6 reports results embedding the forecasts in a regression equation, using them in practice as indicators of future inflation rather than direct forecasts. The tables report the mean forecast error (BIAS), mean absolute error (MAE), root-mean-square error (RMSE) and Theil's U statistic for 1-step-ahead forecasts, the only available. The latter statistics is simply the ratio of the RMSE under test and the RMSE of a benchmark model. In both cases, the benchmark model was a simple regression equation with 12 lags of inflation, together with a constant and seasonal dummies. Tests are also presented according to the terminology of McCracken (2002): his MSE-t, MSE-F tests for forecasting performance, and his ENC-t and ENC-NEW encompassing tests. (Note that the MSE-t test is related to the well-known Diebold-Mariano test.) Since all the tests are based on the MSE of the forecasts (not the RMSE), this latter statistic is also shown with the tests. The standard deviations reported for the RMSE and MSE ratios are based on standard HAC-consistent estimates.

4.2.1 EC Consumer survey forecasts

Both tables 5 and 6 also include an analysis of the importance of biases in the forecasts. In order to do so, forecast MSE are decomposed and the relative importance of each component analysed. The MSE is simply the square of RMSE, its concrete definition being

$$\mathrm{MSE}^{\mathrm{h}} = \mathrm{E}\left(\pi_t - \pi_t^h\right)^2,$$

where π_t^h is the forecast of π_t h periods ahead. Assuming that π_t has mean of $\overline{\pi}_t$ and π_t^h mean of $\overline{\pi}_t^h$, the MSE can be decomposed as

$$\mathrm{MSE}^{\mathrm{h}} = \mathrm{E}\left(\pi_t - \overline{\pi}_t\right)^2 + \mathrm{E}\left(\pi_t^h - \overline{\pi}_t^h\right)^2 - 2\cdot\mathrm{E}\left[\left(\pi_t - \overline{\pi}_t\right)\cdot\left(\pi_t^h - \overline{\pi}_t^h\right)\right] + \left(\overline{\pi}_t - \overline{\pi}_t^h\right)^2.$$

Or, in obvious notation:

$$\mathrm{MSE}^{\mathrm{h}} = \mathrm{var}\left(\pi_t\right) + \mathrm{var}\left(\pi_t^h\right) - 2\cdot\mathrm{cov}\left(\pi_t, \pi_t^h\right) + bias^2.$$

The decomposition is shown in absolute and relative terms. Firstly, the MSE, variances, covariances and biases squared are shown. Then, the RMSE, standard deviations, correlations and biases are shown. Lastly, the MSE, variances, covariances and biases are shown relative to the variance of inflation. This last decomposition shows the relative contribution of each term to the final MSE.

Note that *all* these tests are reported in terms of year-on-year inflation or annual inflation, no matter how the forecasting model was specified. Thus, all numbers shown in the tables are directly comparable.

This decomposition is interesting because it shows the contribution of the bias to forecast errors. Inspecting this is important because of reasons given before about lack of permanent biases in rational forecasts.

Table 5 shows forecasting results for the EC survey forecasts for the next 12 months, taking as forecasts the raw forecast series as derived from last-published inflation, i.e. directly using $\pi_t - \pi_t^h$ as the forecast error. Benchmark forecasts in the table are forecasts based on (16) with 12 lags for inflation and no indicators.[10] In the box, the Theil-U statistic is 3.42 for the indicator-based model, i.e. the raw EC-survey-based forecast is three times as bad as the benchmark. Going to the forecast decomposition, it is clear that the main problem stems from the bias: the MSE is 24% of the variance of inflation, with the bias (squared) contributing 15% of the variance, see last line in the table.

Table 6 takes a different approach. In it, Q6 is used as an indicator in a regression equation like (16), not as raw forecasts. This procedure was adopted because it is not clear whether the quantification done reflects the actual inflation perceived by households. In order to give the measure a fairer chance, it was used as a regressor in an equation: the forecast error is now the difference between actual inflation and the fit of the equation. As before, the benchmark model was an AR(12) model. The regression equation only includes 2 lags for inflation (not 12, as before) and the contemporaneous forecast from Q6—no

[10] The exercise was repeated selecting the lags of the AR process by BIC selection criteria, without significant changes in results.

lags of Q6 were included. This was done both to have non-encompassing models and to simplify the exercise. In any case, the Theil-U statistic is now (compared to the AR(12)), indicating a slightly better performance of the indicator-based model—it is now 0.94. Note that the MSE-t test is highly significant, indicating a statistically better performance of the indicator. Furthermore, the bias is now an insignificant contribution to errors. Lastly, encompassing tests are significant, which implies the presence of useful information in the survey forecasts in addition to information in the alternative forecasts.

In summary, the measure of inflation expectations derived previously is a bad indicator *per se*, but using it in a simple regression beats an alternative AR model. This could indicate that the quantification used might be more successful in capturing the overall shape of household expectations, rather than its precise level.

4.2.2 Consensus Forecasts

For the Consensus Forecasts just one of the two approaches was taken: using the series as raw forecasts. Although the CF forecasts can also be used in the context of forecasting regressions using them as indicators, as was done for the EC survey, it was clear that this was not necessary.

Table 7 reports results for the current- and next-year forecasts. All the definitions of the statistics shown have been reported previously, the only difference in layout between the previous tables and Table 7 being the forecast horizon: 1- to 24-step-ahead forecasts are now reported. The forecasts are grouped according to the horizon with which they were made, assuming that inflation at the time of the forecast was not known. Forecasts are checked against annual inflation (not year-on-year) and are reported assuming that forecasters have not yet observed inflation for the month when the survey was released. Thus, January forecasts are presented as 12-month-ahead forecasts, February ones as 11-month-ahead and so on. Forecasts for the next year are presented as 13- to 24-step-ahead forecasts, i.e. with as many months to the end of forecasted year as there are steps. The benchmark in the table is the naïve no-change forecast, i.e. for each horizon the naïve forecast is the annual inflation rate of the full year previous to the survey. In the case at hand, Consensus beats the benchmark handsomely for all horizons. Table 8 reports an MSE decomposition, as was done for the EC forecasts. Interestingly, the bias doesn't increase much with the horizon: even for the 24-step-ahead forecasts it is only 6% of the variance of inflation while the MSE is close to 51% of the variance, see last line in table 8.

Another interesting comparison can be made using monthly models to build annual inflation forecasts. The procedure, probably not very different from that used by some of the surveyed forecasters, implies running recursive forecasts using an AR(12) model to build monthly forecasts of month-on-month inflation, without any indicator.[11] These monthly forecasts are then used to build annual average forecasts. It is assumed that the monthly models are used to produce forecasts for as many months as

[11] As before, the exercise was also repeated selecting the lags of the AR process by BIC selection criteria, without significant changes in results.

needed to complete the next two years: for instance, if run in April, the forecaster is assumed to build forecasts (using information until March) for April itself and the following 20 months. Observed monthly inflation rates and forecasts are then used to calculate implicit annual inflation rates, which are taken as the benchmark for the Consensus Forecasts. Results against this benchmark are included in Table 9, in which it is evident that short-term forecasts (until roughly 4 months ahead, i.e. for surveys in the period September to December) are considerably worse in the survey than in this alternative benchmark, but long-term ones (from roughly 14 months ahead) are significantly better and actually do beat the benchmark.

Actually, the short-term comparison may not be fully fair to the survey. The procedure chosen for this benchmark has the disadvantage that the AR(12) models are sometimes run on inflation rates that were not considered (or even observed) by the forecasters. For instance, when the AR(12) models are used to produce forecasts in January 1990 they must be run on an aggregation of the inflation rates of the euro area members countries which, at the time, was irrelevant for the forecasters, since the Consensus Forecasts for the period are the weighted average of country forecasts. No forecaster had then in mind an area-wide view when answering the survey. Furthermore, the forecasters could have imperfect information of the last-observed inflation rate at the time, or could have observed yet-to-be-revised data. In this respect, it is useful to remind that the concept itself of HICP did not yet exist during part of the sample. None of these considerations apply to the monthly AR(12) models, which are run directly at the euro area level and with final, fully revised data.

Chart 5 tries to visually assess the point by drawing on the same chart annual average inflation rate, the CF forecasts and the artificial forecasts built using the monthly AR(12) models. Annual average inflation has been calculated as the year-on-year rate of growth of average annual HICP (i.e., the sum of the 12 monthly HICP indices of a year divided by 12). The CF forecasts are raw data: survey respondents are asked to report their forecast for average annual inflation for consumer prices. The AR(12) forecasts were built using the recursively-estimated AR(12) models mentioned previously. In each month, an AR(12) model was estimated using (then) past data and forecasts were built for the rest of the year. AR(12) models were for monthly inflation, but using the initial condition of the HICP index it is possible to map the forecasts (plus available data) into forecasts of the level of the index until the year's end. Once the HICP index is projected for the whole year, it is easy to calculate the annual average inflation rate.

Chart 5 is indicative of departures of what survey forecasters were forecasting from the HICP measure.[12] In the CF, there are sizeable biases between forecasts and actual inflation for the period before 1999, followed by a striking convergence between the two afterwards, basically in the year 1998. Forecasts are afterwards a lot better, with the notable exception of the period around the introduction of euro coins and banknotes, at the end of 2001. Forecasts at this period seriously overshoot inflation, while forecasts

[12] A reminder: HICP as such starts in 1996; it was backdated for this study using national CPI concepts for previous periods. The definition of CPI varied somewhat from country to country, and was slightly different from the HICP concept.

at the beginning of 2002 seriously undershoot it. This behaviour is compatible with forecasters expecting a sharp but short spike of prices at the turn of the year, which actually did not materialise—or at least not with the expected strength. In this sense, professional forecasters and consumers reacted alike to the introduction of the physical euro, although professional forecasters expected a shorter upheaval.

The chart points to a number of exceptionally interesting facts at the period which, unfortunately, can only be analysed when more post-1999 data accrue. The analysis should be repeated for periods in which respondents were directly reporting euro area inflation: but since this information was only collected starting in 1999, as said, the number of observations is clearly insufficient to draw conclusions. The conclusion from this exercise is thus that the CF forecasts are clearly better than the EC forecasts, which may not be surprising, although not overwhelmingly so. As the EC survey, CF forecasts do also convey exploitable information on future inflation, likely for both short and long horizons, although care is needed in doing so.

In summary, all points to relatively good indicator properties of CF forecasts.

5. Conclusion

This paper has analysed rationality in inflation expectations in the euro area. The analysis is non-standard in that two surveys have been used: the European Commission Consumer survey (EC), directed to consumers, and the Consensus Forecasts survey (CF), directed to professional forecasters. The two surveys are without doubt at the extremes of what can be expected of agents in terms of forecasting performance. This fact will be used to highlight the information in the forecasts of one survey relative to the other, rather than in absolute terms.

The very different format of the two surveys has somehow limited the analysis, which has been done only in terms of comparable tests that could be run across surveys. In particular, the consumer survey needed quantifying of the forecasts (which are only qualitative), which implied some care in the analysis of this survey.

Rationality has been tested along three dimensions: i) the extent of absence of bias in the forecasts; ii) the degree to which agents could hope to forecast inflation using their models; and iii) the quality of the forecasts compared to other benchmark (parametric) models.

The conclusions reached are:

- Models used by survey respondents (as they were measured in a so-called forecastability test) were of a sufficient quality to enable them to actually build good forecasts, without a sharp distinction between the two surveys;

- Forecasts themselves seem to be more problematic, in particular the EC survey: bias seems to be present and there is little evidence that past forecasts errors are corrected.

– Survey forecasts do not fare badly compared to simple parametric alternative models used in the paper, the point applying equally well to the EC and CF surveys.

In summary, although the paper has not proved nor disproved that agents in the surveys are rational, at least it can be argued that consumers do not fare that bad compared to professional forecasters. Although the former provide much better forecasts than the latter, as far as the basic ingredients entering rationality are concerned, information in the EC survey should not be ignored.

References

Berk, J. (1999), “Measuring Inflation Expectations: A Survey Data Approach,” *Applied Economics*, 31, 1467-1480.

Berk, J. (2000), “Consumers’ Inflation Expectations and Monetary Policy in Europe,” De Nederlandsche Bank, Satff Reports, No. 55/00.

Carlson, J. and M. Parkin (1975), “Inflation Expectations,” *Economica*, 42, 123-137.

Clements, M.P. and Hendry, D.F. (1999), *Forecasting Economic Time Series*, Cambride University Press, Cambridge.

Diebold, F.X. and R.S. Mariano (1995), “Comparing Predictive Accuracy,” *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X., A.S. Tsay and K.F. Wallis (1997), “Evaluating density forecasts of inflation, the success of professional forecasters,” ESRC Macroeconomic Modelling Bureau Discussion Paper 48, University of Warwick, UK.

Dietrich, J.K. and D.H. Joines (1983), “Rational Expectations, Information Efficiency, and Tests Using Survey Data: A Comment,” *Review of Economics and Statistics*, 65(3), 525-529.

Ehrmann, M. (2006), “Rational Inattention, Inflation Developments And Perceptions After The Euro Area Cash Changeover,” ECB Working Paper Series no. 588.

Figlewski, S. and P. Wachtel (1981), “The Formation of Inflationary Expectations,” *Review of Economics and Statistics*, 63(1), 1-10.

Figlewski, S. and P. Wachtel (1983), “Rational Expectations, Information Efficiency, and Tests Using Survey Data,” *Review of Economics and Statistics*, 65(3), 529-531.

Forsells, M. and G. Kenny (2002), “The rationality of consumers' inflation expectations: survey-based evidence for the euro area”, European Central Bank, Working Paper no. 163.

Giordani, P. and P. Söderlind (2003), "Inflation Forecast Uncertainty," *European Economic Review*, 47, 1037-1059.

Hafer, R.W. and D.H. Resler (1980), “The “Rationality” of Survey-Based Inflation Forecasts,” *Review*, Saint Louis Fed, November, 3-11.

Hendry, D. F. and M. P. Clements (2001), “Pooling of Forecasts,” *Econometrics Journal*, 5, 1-26.

Łyziak, T. and E. Stanisławska (2006), “Consumer inflation expectations. Survey Questions and quantification methods,” National Bank of Poland’s Materiały i Studia Paper no. 37.

Mankiw, N.G. and R. Reis (2003), "Disagreement about inflation expectations", NBER working paper no. 9796.

McCracken, M.W and E.C. Todd (2002), "Evaluating Long-Horizon Forecasts", Federal Reserve Bank of Kansas City, working paper RWP 01-14.

Monthly Bulletin of the ECB, July 2002, p. 18-19.

Monthly Bulletin of the ECB, October 2003, p. 24-25.

Monthly Bulletin of the ECB, April 2005, p. 30-32.

Monthly Bulletin of the ECB, July 2006, p. 59-67.

Monthly Bulletin of the ECB, September 2006, p. 52-53.

Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements,", *Econometrica*, 21, 315-335.

Nielsen, H. (2003) "Inflation Expectations in the EU—Results from Survey Data," FU and HU Berlin, mimeo.

Pesaran, M. (1987), *The Limits to Rational Expectations*, Basil Blackwell, Oxford.

Table 1. EC Household Survey Forecast, Next 12 Months: dynamic REH analysis (I)

Variable: INFLATION

Sampling:

1986:10 to ...	2003:08	2002:08	2001:08	2000:08	1999:08	1998:08	1997:08
Constant	0.001	0.001	0.001	0.001	0.001	0.001	0.002
t-stat	2.394	2.542	2.577	2.554	1.913	1.853	2.514
t-stat (stability)		-0.193	-0.257	-0.314	0.035	-0.463	-2.305
INFLATION[-1]	1.100	1.093	1.120	1.109	1.100	1.088	1.088
t-stat	16.382	15.926	15.789	15.106	14.380	13.571	12.981
t-stat (stability)		0.115	-0.285	-0.121	0.007	0.188	0.179
INFLATION[-2]	-0.322	-0.307	-0.342	-0.316	-0.298	-0.289	-0.297
t-stat	-3.227	-3.011	-3.208	-2.866	-2.597	-2.419	-2.375
t-stat (stability)		-0.148	0.198	-0.066	-0.244	-0.332	-0.249
INFLATION[-3]	0.178	0.168	0.181	0.179	0.169	0.165	0.161
t-stat	2.518	2.333	2.380	2.282	2.066	1.932	1.789
t-stat (stability)		0.132	-0.047	-0.027	0.117	0.172	0.230
INFLATION[-8]	0.206	0.225	0.207	0.185	0.188	0.193	0.186
t-stat	4.517	4.791	4.127	3.679	3.608	3.503	3.196
t-stat (stability)		-0.420	-0.027	0.455	0.375	0.268	0.430
INFLATION[-11]	-0.218	-0.241	-0.224	-0.216	-0.214	-0.214	-0.210
t-stat	-5.132	-5.559	-5.012	-4.834	-4.593	-4.402	-4.171
t-stat (stability)		0.550	0.145	-0.049	-0.097	-0.100	-0.189
ECQ6[-18]	-0.016	-0.021	-0.021	-0.003	-0.008	-0.012	-0.001
t-stat	-0.250	-0.323	-0.320	-0.041	-0.114	-0.169	-0.021
t-stat (stability)		0.080	0.079	-0.209	-0.131	-0.066	-0.227
ECQ6[-19]	0.133	0.148	0.150	0.134	0.135	0.134	0.127
t-stat	1.856	2.034	2.067	1.863	1.832	1.755	1.631
t-stat (stability)		-0.202	-0.239	-0.011	-0.022	-0.009	0.088
ECQ6[-21]	-0.089	-0.093	-0.100	-0.103	-0.099	-0.098	-0.106
t-stat	-2.419	-2.449	-2.649	-2.701	-2.551	-2.387	-2.522
t-stat (stability)		0.100	0.311	0.372	0.284	0.247	0.472
Decreasing N Chow-test		1.186	1.297	1.381	1.150	0.915	0.865
Significance		0.296	0.173	0.092	0.262	0.646	0.747

Degrees of Freedom	194
Std Error of Dependent Variable	0.035306
Standard Error of Estimate	0.001760
Sum of Squared Residuals	0.000601
R**2	0.9706
R**2 BAR	0.9694
Durbin-Watson	2.0415
Centered F test: F(8,194)	799.9621 (0.00)
Uncentered LM test: Chi(9)	202.2492 (0.00)

```
L-M autocorrelation test (lag 1):
  F(1,192) =       1.11322 ; Signif:       0.29271
L-M autocorrelation test (lags 1 to 2):
  F(2,190) =       3.66200 ; Signif:       0.02751
L-M autocorrelation test (lags 1 to 12):
  F(12,170) =       5.34316 ; Signif:   1.22132e-07
L-M autocorrelation test (lags 1 to 24):
  F(24,146) =       3.96190 ; Signif:   1.07759e-07
Normality test:
  Chi(2) =      17.66357 ; Signif:   1.46017e-04
ARCH test:
  Chi(1) =       0.86141 ; Signif:       0.35335
RESET test:
  F(1,193) =       0.08595 ; Signif:       0.76971
```

Expectations based on previous-month inflation.

Table 1 (cont.). EC Household Survey Forecast, Next 12 Months: dynamic REH analysis (I)

```
Long-Run Analysis of series INFLATION

Sum of dynamic parameters

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant             9.11156e-04   3.80546e-04       2.39433
INFLATION               -0.05670       0.02250      -2.52012 (ECM=0)
ECQ6                     0.02849       0.02070       1.37624

Static Long-Run Solution

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant                 0.01607       0.00608       2.64240
ECQ6                     0.50235       0.23671       2.12224

Tests of the Significance of each variable

Constant           F(1,194)=       5.73284 ; Signif.=       0.01760
INFLATION          F(5,194)=     870.79541 ; Signif.=  8.81360e-131
ECQ6               F(3,194)=       2.51943 ; Signif.=       0.05928

Tests of the Significance of each lag

Lag 0              F(1,194)=       5.73284 ; Signif.=       0.01760
Lag 1              F(1,194)=     268.36024 ; Signif.=   1.93948e-38
Lag 2              F(1,194)=      10.41598 ; Signif.=       0.00147
Lag 3              F(1,194)=       6.33967 ; Signif.=       0.01262
Lag 8              F(1,194)=      20.40243 ; Signif.=   1.08829e-05
Lag 11             F(1,194)=      26.33893 ; Signif.=   6.92815e-07
Lag 18             F(1,194)=       0.06262 ; Signif.=       0.80267
Lag 19             F(1,194)=       3.44648 ; Signif.=       0.06490
Lag 21             F(1,194)=       5.84957 ; Signif.=       0.01650
```

Expectations based on previous-month inflation.

Table 2. EC Household Survey Forecast, Next 12 Months: dynamic forecast analysis (II)

```
Long-Run Analysis of series ECQ6

Sum of dynamic parameters

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant                -0.00188   3.79388e-04      -4.96177
INFLATION                0.34036       0.03405       9.99701
ECQ6                    -0.32273       0.03095     -10.42638 (ECM=0)
SEASONAL             6.07084e-04   5.62391e-04       1.07947

Static Long-Run Solution

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant                -0.00583       0.00108      -5.39819
INFLATION                1.05460       0.03741      28.18917
SEASONAL                 0.00188       0.00175       1.07416

Tests of the Significance of each variable

Constant           F(1,185)=      24.61912 ; Signif.=   1.57720e-06
BCH                F(6,185)=      33.41480 ; Signif.=   3.91954e-27
ECQ6               F(7,185)=      86.99230 ; Signif.=   4.09559e-55
SEASONAL           F(2,185)=       2.94381 ; Signif.=       0.05514

Tests of the Significance of each lag

Lag 0              F(1,185)=      24.61912 ; Signif.=   1.57720e-06
Lag 1              F(3,185)=     115.99419 ; Signif.=   2.74388e-42
Lag 2              F(1,185)=      41.25502 ; Signif.=   1.09779e-09
Lag 3              F(1,185)=       8.48118 ; Signif.=       0.00403
Lag 4              F(2,185)=       2.46783 ; Signif.=       0.08756
Lag 5              F(1,185)=       5.03248 ; Signif.=       0.02606
Lag 6              F(1,185)=       2.62671 ; Signif.=       0.10678
Lag 8              F(1,185)=       2.16165 ; Signif.=       0.14319
Lag 9              F(1,185)=       1.85553 ; Signif.=       0.17480
Lag 11             F(2,185)=      11.03435 ; Signif.=   2.97028e-05
Lag 12             F(2,185)=      29.01871 ; Signif.=   1.09299e-11
```

Expectations based on previous-month inflation.

Table 3. EC Household Survey Forecast, Next 12 Months: dynamic REH analysis (I)

```
Long-Run Analysis of serie INFLATION

Sum of dynamic parameters

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant                 0.00103   5.04200e-04       2.03865
INFLATION               -0.02369       0.01751      -1.35269 (ECM=0)
ECQ6B                   -0.01775       0.03002      -0.59125

Static Long-Run Solution

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
Constant                 0.04339       0.03736       1.16132
ECQ6B                   -0.74918       1.68837      -0.44373

Tests of the Significance of each variable

Constant           F(1,205)=       4.15609 ; Signif.=       0.04277
INFLATION          F(3,205)=    1521.19975 ; Signif.=  9.32533e-140
ECQ6B              F(2,205)=      11.84084 ; Signif.=   1.36005e-05

Tests of the Significance of each lag

Lag 0              F(1,205)=       4.15609 ; Signif.=       0.04277
Lag 1              F(1,205)=    1829.01124 ; Signif.=  4.13923e-104
Lag 8              F(1,205)=      12.30267 ; Signif.=   5.55739e-04
Lag 11             F(1,205)=      26.73899 ; Signif.=   5.52074e-07
Lag 12             F(1,205)=      23.54611 ; Signif.=   2.41182e-06
Lag 13             F(1,205)=      21.87781 ; Signif.=   5.26848e-06
```

Expectations based on perceived inflation.

Table 4. EC Household Survey Forecast, Next 12 Months: dynamic forecast analysis (II)

```
Long-Run Analysis of serie ECQ6B (Forecasts relative to Perceived Infl.)

Sum of dynamic parameters

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
INFLATION                0.04187       0.01521       2.75333
ECQ6B                   -0.05113       0.01890      -2.70588 (ECM=0)
SEASONAL             2.37815e-04   3.48130e-04       0.68312

Static Long-Run Solution

                       Parameter Standard Err.   t-statistic
                   ------------- ------------- -------------
INFLATION                0.81889       0.07013      11.67692
SEASONAL                 0.00465       0.00693       0.67073

Tests of the Significance of each variable

INFLATION          F(8,184)=       9.08611 ; Signif.=   1.70237e-10
ECQ6B              F(7,184)=     366.52654 ; Signif.=  1.91938e-104
SEASONAL           F(1,184)=       0.46666 ; Signif.=       0.49539

Tests of the Significance of each lag

Lag 0              F(1,184)=      16.23632 ; Signif.=   8.17418e-05
Lag 1              F(2,184)=     116.37064 ; Signif.=   2.16531e-33
Lag 2              F(1,184)=      35.72137 ; Signif.=   1.15639e-08
Lag 3              F(1,184)=      16.17984 ; Signif.=   8.40113e-05
Lag 5              F(2,184)=       5.93671 ; Signif.=       0.00317
Lag 6              F(2,184)=       6.20813 ; Signif.=       0.00246
Lag 7              F(2,184)=       2.41653 ; Signif.=       0.09206
Lag 8              F(1,184)=       1.33951 ; Signif.=       0.24862
Lag 11             F(2,184)=       2.93681 ; Signif.=       0.05553
Lag 12             F(2,184)=       2.37490 ; Signif.=       0.09587
```

Expectations based on perceived inflation.

Table 5. Year-on-year inflation forecasts for 1985M1 to 2003M8, EC Consumer Survey Q6

```
Forecast Statistics for Series ANNUALISED INFLATION
Step  Mean Error Mean Abs Error  RMS Error      Theil U  N.Obs
   1  0.004167557  0.004750760  0.005273542      3.42316  223

    ==== Candidate ==== ==== Benchmark ====
Step N.Obs.        RMSE N.Obs.        RMSE        Ratio     Std.Dev.
   1    223    0.0052735    223    0.0015405    3.4231625    0.2426682
    ==== Candidate ==== ==== Benchmark ====
Step N.Obs.         MSE N.Obs.         MSE        Ratio     Std.Dev.        MSE-t        MSE-F        ENC-t      ENC-NEW
   1    223    0.0000278    223    0.0000024   11.7180415    1.6613856 -235.9782689 -203.9695161   46.5902141   13.5701224


Forecast MSE Decomposition for ANNUALISED INFLATION

Actual numbers

Step             MSE           var(x)           var(f)         cov(x,f)          bias^2  Obs.
   1     0.000027810      0.000115462      0.000134109      0.000119565     0.000017369   223

Decomposition in terms of RMSE, standard deviations, bias and correlations

Step            RMSE           std(x)           std(f)        corr(x,f)            bias  Obs.
   1     0.005273542      0.010745329      0.011580562      0.960847264     0.004167557   223

Decomposition relative to var(x)

Step             MSE           var(x)           var(f)         cov(x,f)          bias^2  Obs.
   1     0.240860350      1.000000000      1.161501555      1.035533745     0.150426285   223
```

Expectations based on actual past inflation. Benchmark model in the table: an AR(12) model with no indicators.

Table 6. Year-on-year inflation forecasts for 1990M1 to 2003M8, Consumer Survey Q6 used as indicator

```
Forecast Statistics for Series ANNUALISED INFLATION
Step  Mean Error Mean Abs Error  RMS Error      Theil U  N.Obs
   1 -0.000202871  0.001000051  0.001374037      0.93712  162

    ==== Candidate ==== ==== Benchmark ====
Step N.Obs.        RMSE N.Obs.        RMSE       Ratio     Std.Dev.
   1    162    0.0013740    162    0.0014662    0.9371195    0.0460773
    ==== Candidate ==== ==== Benchmark ====
Step N.Obs.         MSE N.Obs.         MSE       Ratio     Std.Dev.       MSE-t        MSE-F        ENC-t      ENC-NEW
   1    162    0.0000019    162    0.0000021    0.8781930    0.0863599   16.1365059   22.4696919   34.5289196   27.7337023


Forecast MSE Decomposition for ANNUALISED INFLATION

Actual numbers

Step              MSE            var(x)            var(f)          cov(x,f)           bias^2  Obs.
   1      0.000001888       0.000100909       0.000101406       0.000100234      0.000000041   162

Decomposition in terms of RMSE, standard deviations, bias and correlations

Step             RMSE            std(x)            std(f)         corr(x,f)             bias  Obs.
   1      0.001374037       0.010045331       0.010070038       0.990874523     -0.000202871   162

Decomposition relative to var(x)

Step              MSE            var(x)            var(f)          cov(x,f)           bias^2  Obs.
   1      0.018709754       1.000000000       1.004925232       0.993311670      0.000407862   162
```

Expectations based on actual past inflation. Benchmark model in the table: an AR(12) model with no indicators.

Table 7. Year-on-year inflation forecasts for 1990M1 to 2003M8, Consensus Forecasts vs naïve randon-walk model

Forecast Statistics for Series ANNUALISED INFLATION

Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-0.000683861	0.001353355	0.001774527	0.37797	13
2	-0.000813253	0.001590849	0.001963564	0.41823	13
3	-0.001029085	0.001704120	0.002147715	0.45745	13
4	-0.000986183	0.001907358	0.002434553	0.51855	13
5	-0.000786156	0.002157839	0.002737028	0.58298	13
6	-0.000861111	0.002380377	0.002877794	0.61296	13
7	-0.000776157	0.002379007	0.002793064	0.59491	13
8	-0.000587648	0.002580594	0.003010594	0.64124	13
9	-0.000429618	0.002900693	0.003183666	0.67811	13
10	-0.000261930	0.003310909	0.003691132	0.78620	13
11	-0.000539447	0.003375603	0.003804732	0.81039	13
12	-0.000935987	0.003687463	0.004220636	0.89898	13
13	-0.001617971	0.004347062	0.004731039	0.57854	12
14	-0.001968359	0.004530783	0.004818340	0.58922	12
15	-0.002176382	0.004572139	0.004941977	0.60434	12
16	-0.002184737	0.004671104	0.005202193	0.63616	12
17	-0.001913182	0.005379901	0.005713723	0.69871	12
18	-0.002029345	0.005465735	0.005834233	0.71345	12
19	-0.002074932	0.005460021	0.005900761	0.72159	12
20	-0.002205500	0.005587552	0.006054259	0.74036	12
21	-0.002077003	0.005843517	0.006408564	0.78368	12
22	-0.002136942	0.005892881	0.006476279	0.79196	12
23	-0.002022551	0.006075831	0.006620040	0.80954	12
24	-0.002287526	0.006088240	0.006628926	0.81063	12

	==== Candidate ====		==== Benchmark ====							
Step	N.Obs.	MSE	N.Obs.	MSE	Ratio	Std.Dev.	MSE-t	MSE-F	ENC-t	ENC-NEW
1	13	0.0000031	13	0.0000220	0.1428589	0.0717161	9.5119705	77.9988881	8.7554961	76.2292951
2	13	0.0000039	13	0.0000220	0.1749169	0.0817502	10.9653534	61.3209896	9.2948961	59.8540598
3	13	0.0000046	13	0.0000220	0.2092644	0.0867707	12.6476123	49.1223798	9.8949484	49.1979155
4	13	0.0000059	13	0.0000220	0.2688934	0.1206842	10.5059259	35.3462910	9.5323461	35.7424555
5	13	0.0000075	13	0.0000220	0.3398602	0.1701777	8.5309780	25.2510165	8.8759344	24.8056410
6	13	0.0000083	13	0.0000220	0.3757175	0.1562080	9.1681200	21.6004682	9.4456145	21.3480274
7	13	0.0000078	13	0.0000220	0.3539189	0.1300540	10.4940261	23.7315812	10.8149927	20.0884468
8	13	0.0000091	13	0.0000220	0.4111934	0.0988714	11.9116121	18.6152908	13.0041203	15.0664251
9	13	0.0000101	13	0.0000220	0.4598296	0.0783873	12.6814626	15.2713427	13.6447929	12.8134299
10	13	0.0000136	13	0.0000220	0.6181032	0.1299324	7.8022043	8.0320873	10.6257169	8.0278888
11	13	0.0000145	13	0.0000220	0.6567348	0.0856618	10.9724679	6.7948995	13.5948378	7.3510185
12	13	0.0000178	13	0.0000220	0.8081606	0.1390074	4.9596011	3.0859107	17.8293409	5.1790333
13	12	0.0000224	12	0.0000669	0.3347137	0.0318141	37.1962850	23.8515322	44.1852711	18.4195423
14	12	0.0000232	12	0.0000669	0.3471805	0.0235527	37.5184891	22.5641557	44.1701447	17.0620468
15	12	0.0000244	12	0.0000669	0.3652262	0.0231983	35.6989583	20.8563539	38.5659637	15.4920549
16	12	0.0000271	12	0.0000669	0.4047001	0.0313345	40.6697706	17.6515889	45.5482998	12.4904570
17	12	0.0000326	12	0.0000669	0.4882012	0.0280244	36.3182087	12.5800314	41.3172398	8.8437033
18	12	0.0000340	12	0.0000669	0.5090118	0.0320690	38.1001630	11.5750908	43.5735009	8.2400082
19	12	0.0000348	12	0.0000669	0.5206867	0.0380488	38.1035243	11.0464886	44.3569113	7.9862631
20	12	0.0000367	12	0.0000669	0.5481285	0.0449132	35.6745300	9.8926784	40.6891563	7.4008524
21	12	0.0000411	12	0.0000669	0.6141604	0.0582343	26.7458998	7.5388689	32.7432985	6.0118458
22	12	0.0000419	12	0.0000669	0.6272079	0.0613307	23.5627606	7.1324121	29.1954277	5.7804767
23	12	0.0000438	12	0.0000669	0.6553626	0.0543518	23.6428364	6.3104753	31.6762938	5.3391964
24	12	0.0000439	12	0.0000669	0.6571231	0.0635708	21.4383275	6.2614189	31.1533009	5.1975816

Table 8. Year-on-year inflation forecasts decomposition for 1990M1 to 2003M8, Consensus Forecasts

Forecast MSE Decomposition for ANNUALISED INFLATION

Decomposition in terms of RMSE, standard deviations, bias and correlations

Step	RMSE	std(x)	std(f)	corr(x,f)	bias	Obs.
1	0.001774527	0.010252551	0.010254941	0.987248939	-0.000683861	13
2	0.001963564	0.010252551	0.010235820	0.984782626	-0.000813253	13
3	0.002147715	0.010252551	0.010172084	0.982993595	-0.001029085	13
4	0.002434553	0.010252551	0.010292039	0.976530771	-0.000986183	13
5	0.002737028	0.010252551	0.010111529	0.966945778	-0.000786156	13
6	0.002877794	0.010252551	0.010186719	0.963922513	-0.000861111	13
7	0.002793064	0.010252551	0.010140466	0.965439420	-0.000776157	13
8	0.003010594	0.010252551	0.010205971	0.958350508	-0.000587648	13
9	0.003183666	0.010252551	0.010312453	0.952957204	-0.000429618	13
10	0.003691132	0.010252551	0.010120218	0.934760076	-0.000261930	13
11	0.003804732	0.010252551	0.010290518	0.932782004	-0.000539447	13
12	0.004220636	0.010252551	0.010647517	0.923135524	-0.000935987	13
13	0.004731039	0.009287529	0.010542287	0.907107789	-0.001617971	12
14	0.004818340	0.009287529	0.010388145	0.906039742	-0.001968359	12
15	0.004941977	0.009287529	0.010258680	0.901638310	-0.002176382	12
16	0.005202193	0.009287529	0.009626797	0.875993504	-0.002184737	12
17	0.005713723	0.009287529	0.009213109	0.830654198	-0.001913182	12
18	0.005834233	0.009287529	0.009093485	0.823089104	-0.002029345	12
19	0.005900761	0.009287529	0.008987131	0.817754619	-0.002074932	12
20	0.006054259	0.009287529	0.008774487	0.806569329	-0.002205500	12
21	0.006408564	0.009287529	0.008762377	0.775868823	-0.002077003	12
22	0.006476279	0.009287529	0.008393980	0.765408191	-0.002136942	12
23	0.006620040	0.009287529	0.008454731	0.751407940	-0.002022551	12
24	0.006628926	0.009287529	0.008604660	0.760726509	-0.002287526	12

Decomposition relative to var(x)

Step	MSE	var(x)	var(f)	cov(x,f)	bias^2	Obs.
1	0.029957221	1.000000000	1.000466216	0.987479048	0.004449101	13
2	0.036679730	1.000000000	0.996738913	0.983175584	0.006291985	13
3	0.043882315	1.000000000	0.984364550	0.975278546	0.010074856	13
4	0.056386411	1.000000000	1.007717777	0.980291851	0.009252336	13
5	0.071268010	1.000000000	0.972679527	0.953645599	0.005879681	13
6	0.078787195	1.000000000	0.987199046	0.957733077	0.007054303	13
7	0.074216076	1.000000000	0.978254639	0.954884812	0.005731062	13
8	0.086226435	1.000000000	0.990934000	0.953996415	0.003285264	13
9	0.096425340	1.000000000	1.011719434	0.958524998	0.001755902	13
10	0.129614992	1.000000000	0.974351993	0.922694845	0.000652689	13
11	0.137715972	1.000000000	1.007419944	0.936236204	0.002768436	13
12	0.169469661	1.000000000	1.078531391	0.958698083	0.008334435	13
13	0.259485308	1.000000000	1.288455089	1.029659282	0.030348783	12
14	0.269150097	1.000000000	1.251052752	1.013409705	0.044916756	12
15	0.283139934	1.000000000	1.220063872	0.995918133	0.054912328	12
16	0.313741901	1.000000000	1.074393146	0.907992999	0.055334754	12
17	0.378475752	1.000000000	0.984038365	0.823998231	0.042433850	12
18	0.394609122	1.000000000	0.958650590	0.805892334	0.047743200	12
19	0.403660012	1.000000000	0.936357754	0.791305002	0.049912263	12
20	0.424934112	1.000000000	0.892571633	0.762014514	0.056391507	12
21	0.476125097	1.000000000	0.890109694	0.731998290	0.050011982	12
22	0.486240095	1.000000000	0.816837231	0.691768636	0.052940136	12
23	0.508066869	1.000000000	0.828703584	0.684030376	0.047424038	12
24	0.509431709	1.000000000	0.858355188	0.704793789	0.060664100	12

Table 9. Year-on-year inflation forecasts for 1990M1 to 2003M8, Consensus Forecasts vs monthly AR(12) model

Forecast Statistics for Series ANNUALISED INFLATION

Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-0.000683861	0.001353355	0.001774527	17.49043	13
2	-0.000813253	0.001590849	0.001963564	5.53522	13
3	-0.001029085	0.001704120	0.002147715	4.92680	13
4	-0.000986183	0.001907358	0.002434553	3.12130	13
5	-0.000786156	0.002157839	0.002737028	1.99215	13
6	-0.000861111	0.002380377	0.002877794	1.96522	13
7	-0.000776157	0.002379007	0.002793064	1.34046	13
8	-0.000587648	0.002580594	0.003010594	1.47360	13
9	-0.000429618	0.002900693	0.003183666	1.55435	13
10	-0.000261930	0.003310909	0.003691132	1.64460	13
11	-0.000539447	0.003375603	0.003804732	1.01949	13
12	-0.000935987	0.003687463	0.004220636	1.05219	13
13	-0.001617971	0.004347062	0.004731039	0.88773	12
14	-0.001968359	0.004530783	0.004818340	0.68981	12
15	-0.002176382	0.004572139	0.004941977	0.65546	12
16	-0.002184737	0.004671104	0.005202193	0.56853	12
17	-0.001913182	0.005379901	0.005713723	0.56832	12
18	-0.002029345	0.005465735	0.005834233	0.70995	12
19	-0.002074932	0.005460021	0.005900761	0.74420	12
20	-0.002205500	0.005587552	0.006054259	0.84923	12
21	-0.002077003	0.005843517	0.006408564	0.81216	12
22	-0.002136942	0.005892881	0.006476279	0.71578	12
23	-0.002022551	0.006075831	0.006620040	0.68857	12
24	-0.002287526	0.006088240	0.006628926	0.62100	12

	==== Candidate ====		==== Benchmark ====							
Step	N.Obs.	MSE	N.Obs.	MSE	Ratio	Std.Dev.	MSE-t	MSE-F	ENC-t	ENC-NEW
1	13	0.0000031	13	0.0000000	305.9150145	163.6084617	-8.5517656	-12.9575045	3.6474527	0.2132817
2	13	0.0000039	13	0.0000001	30.6386187	23.7209616	-10.0053757	-12.5756989	3.9070200	1.1292705
3	13	0.0000046	13	0.0000002	24.2733814	11.0751167	-11.1364347	-12.4644339	0.7496508	0.2082628
4	13	0.0000059	13	0.0000006	9.7424888	2.4172054	-10.7265918	-11.6656387	-5.3240001	-0.9489707
5	13	0.0000075	13	0.0000019	3.9686605	1.7798121	-11.0390875	-9.7243355	0.8459432	0.2160254
6	13	0.0000083	13	0.0000021	3.8620795	1.5587404	-8.7527599	-9.6339378	-5.4030380	-1.7418451
7	13	0.0000078	13	0.0000043	1.7968277	0.8735592	-4.3317273	-5.7650271	5.5532974	2.6498185
8	13	0.0000091	13	0.0000042	2.1714866	0.3664431	-15.8736763	-7.0133179	10.6441599	2.4996018
9	13	0.0000101	13	0.0000042	2.4159968	0.5778386	-18.6001116	-7.6191981	12.8073107	2.6797373
10	13	0.0000136	13	0.0000050	2.7047195	0.7303603	-15.0278733	-8.1935865	7.1990153	2.4321765
11	13	0.0000145	13	0.0000139	1.0393564	0.2263533	-0.6595921	-0.4922590	15.7004742	10.2662570
12	13	0.0000178	13	0.0000161	1.1070997	0.1738772	-2.1861939	-1.2576068	9.2160527	5.8753216
13	12	0.0000224	12	0.0000284	0.7880729	0.3295677	1.8665240	3.2270183	6.5072472	6.3568154
14	12	0.0000232	12	0.0000488	0.4758314	0.2837439	3.3211308	13.2190192	5.2015906	15.8217100
15	12	0.0000244	12	0.0000568	0.4296341	0.1615080	5.9485628	15.9307419	9.5699939	19.6231230
16	12	0.0000271	12	0.0000837	0.3232301	0.1167692	7.9273257	25.1252545	10.4751706	23.5080209
17	12	0.0000326	12	0.0001011	0.3229839	0.1584713	5.4068872	25.1535551	7.3686037	26.1423777
18	12	0.0000340	12	0.0000675	0.5040250	0.0744040	16.8081641	11.8083446	15.4798859	13.8759289
19	12	0.0000348	12	0.0000629	0.5538334	0.0752564	11.4138859	9.6671674	9.4441352	11.4036450
20	12	0.0000367	12	0.0000508	0.7211992	0.1096044	7.4490583	4.6389543	27.7301082	7.0039496
21	12	0.0000411	12	0.0000623	0.6595965	0.1285099	7.1491835	6.1929393	26.7220337	8.4130367
22	12	0.0000419	12	0.0000819	0.5123386	0.1159006	10.2509814	11.4220101	12.8140128	12.5172305
23	12	0.0000438	12	0.0000924	0.4741340	0.0605514	25.3167045	13.3093016	18.4719565	14.4430462
24	12	0.0000439	12	0.0001139	0.3856463	0.0588220	26.3456188	19.1165948	47.8365519	18.1394710

Chart 1

Static Rationality Tests, Consensus Forecasts

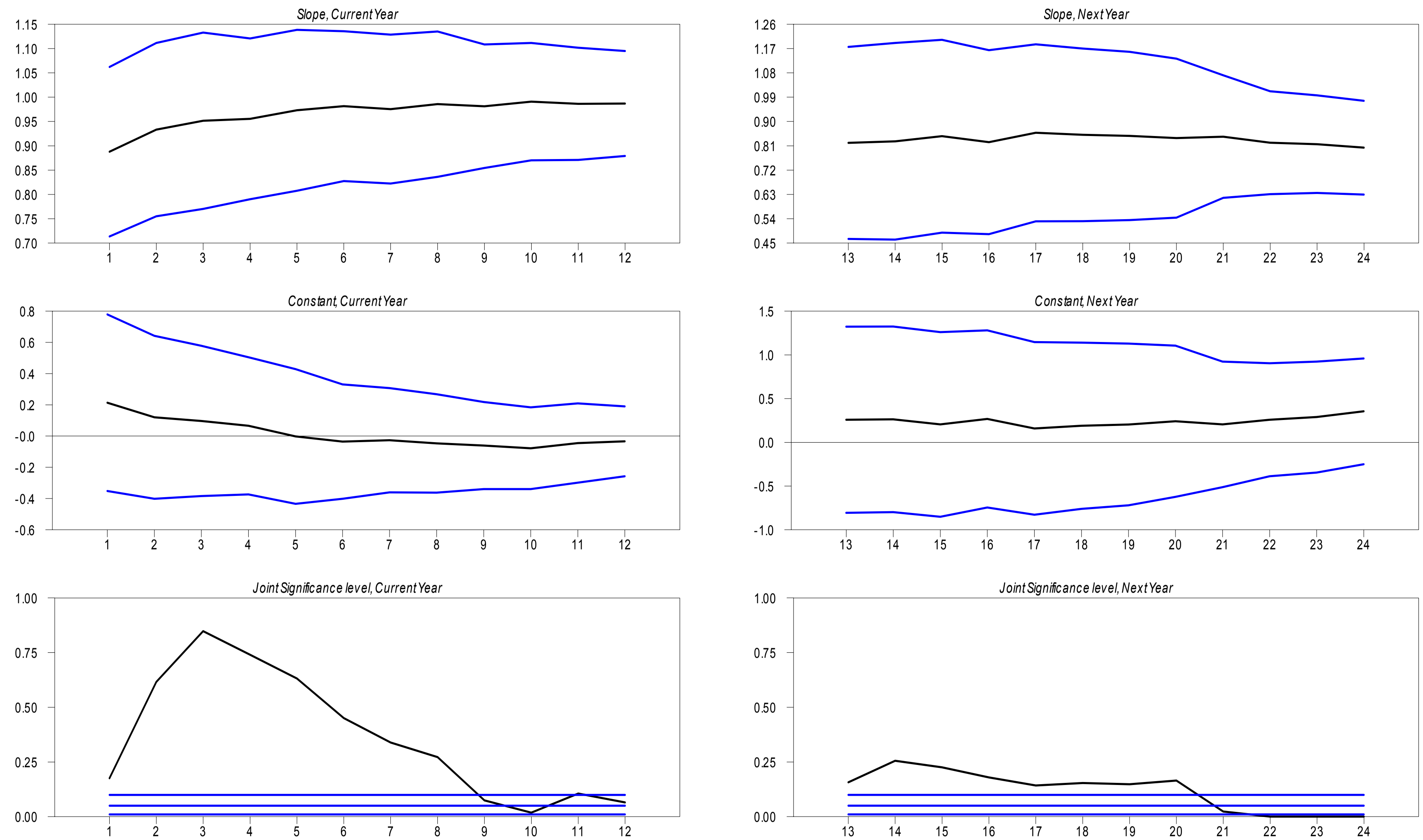


Chart 2

Dynamic Rationality Tests, Consensus Forecasts

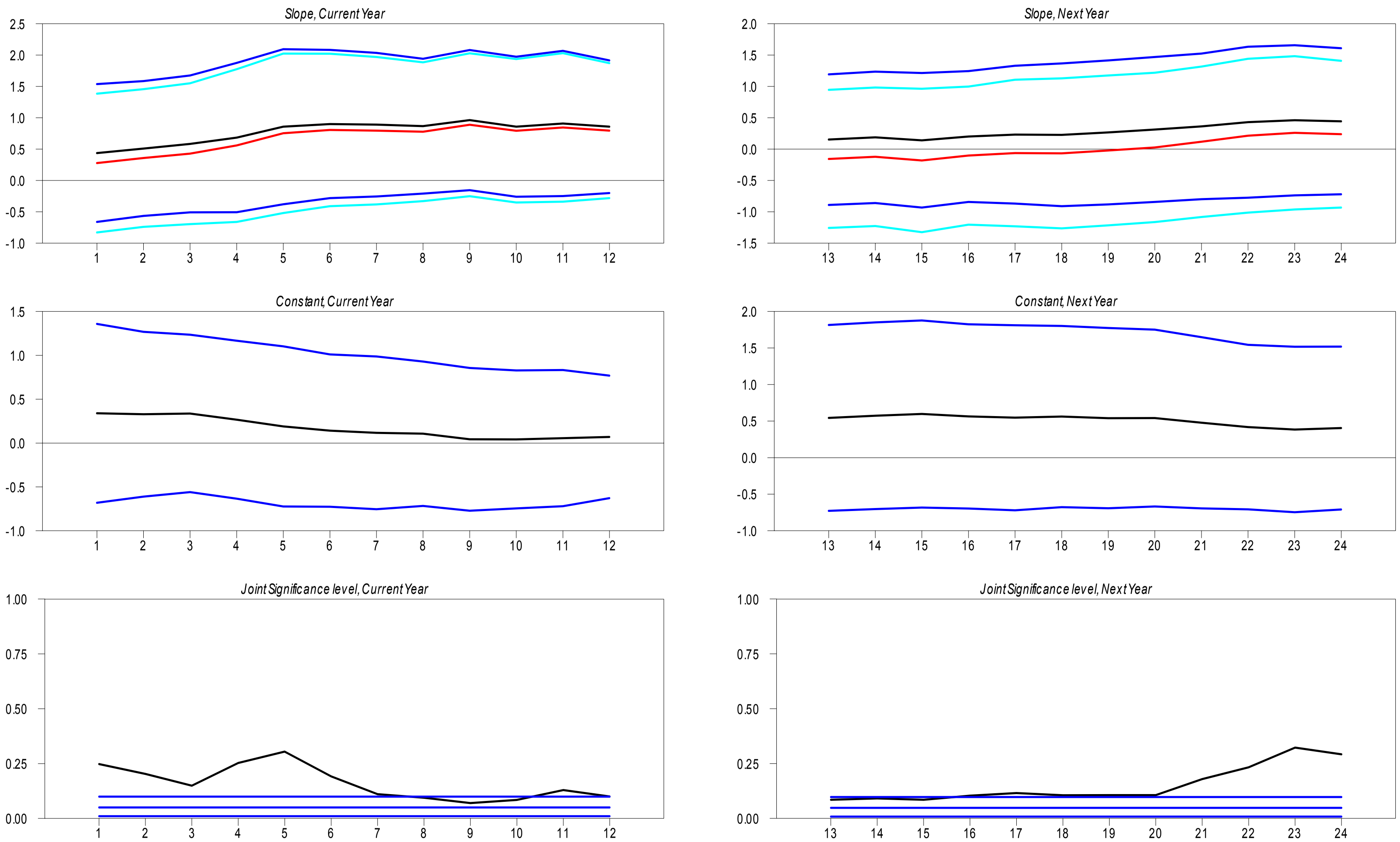


Chart 3

Static Forecastability Tests, Consensus Forecasts

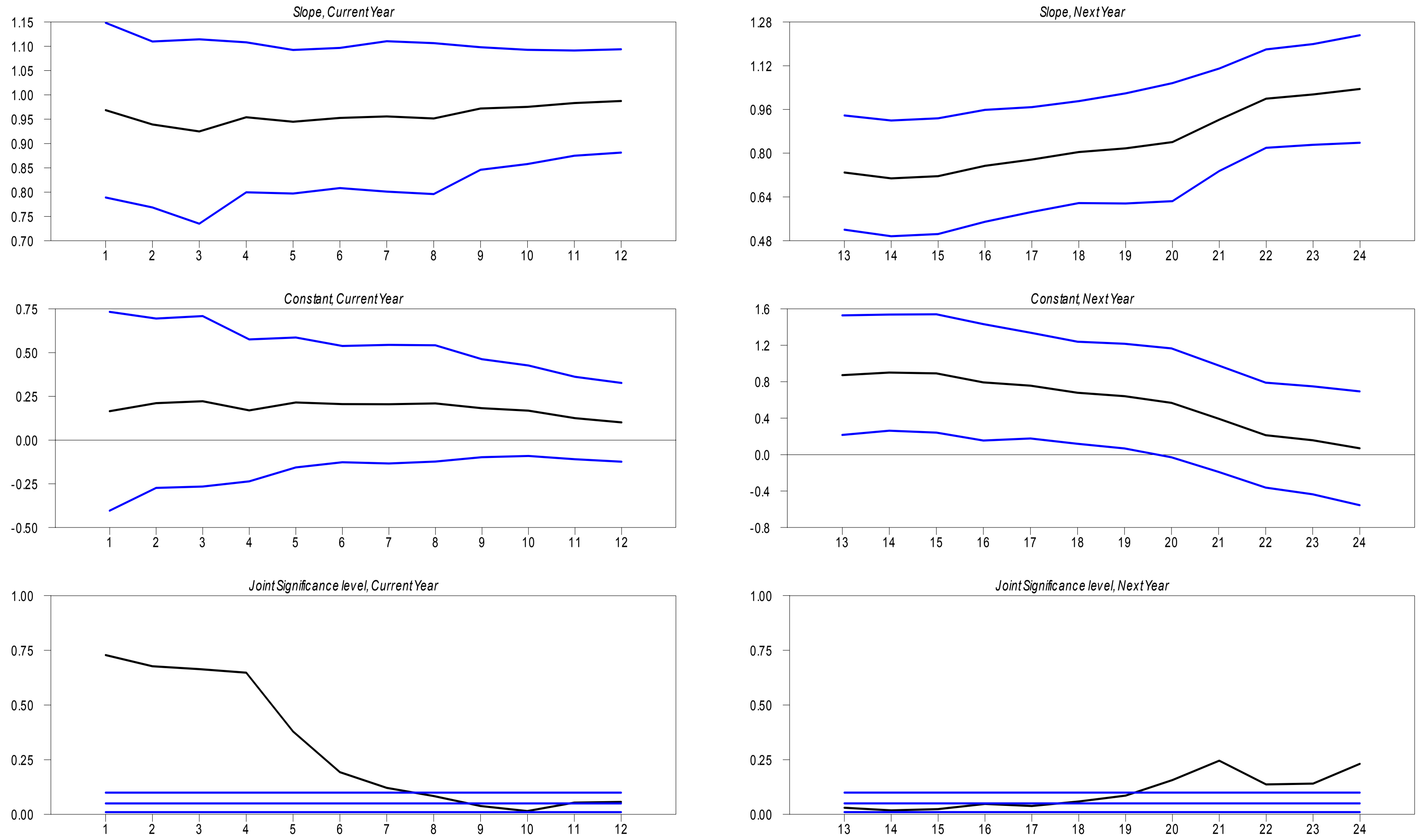


Chart 4

Dynamic Forecastability Tests, Consensus Forecasts

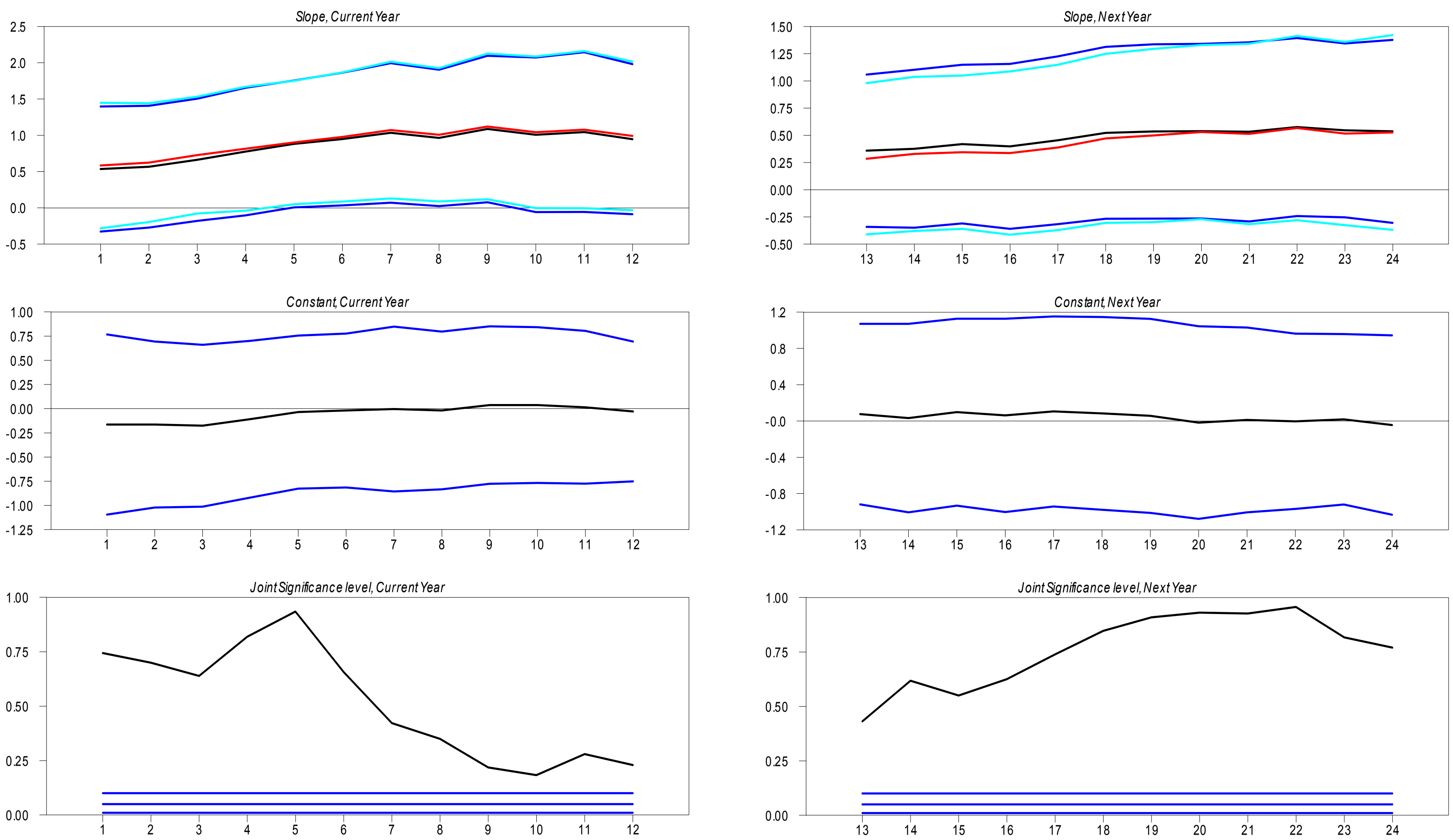


Chart 5

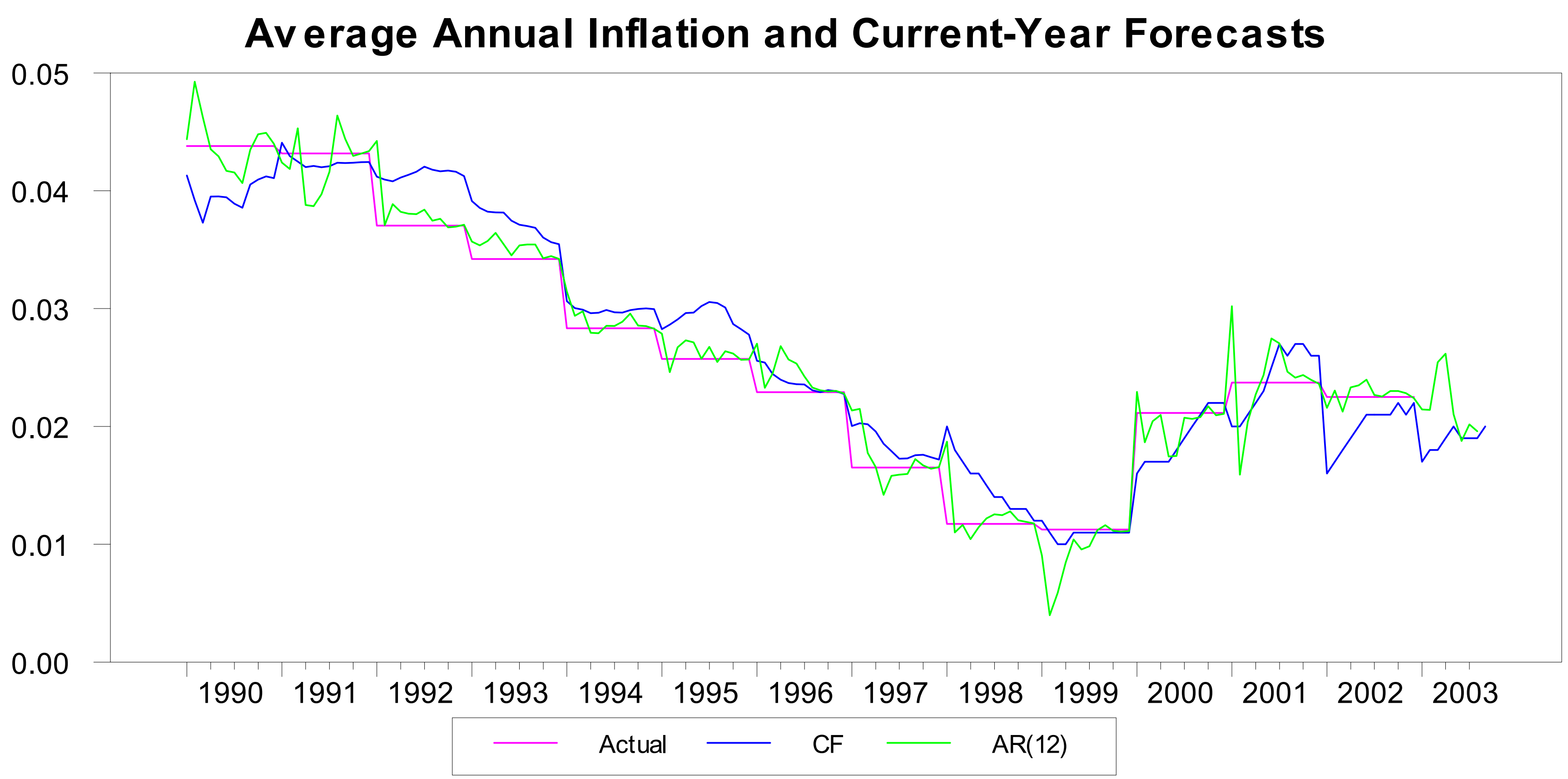


Chart shows average annual inflation, as measured by the rate of growth of average annual HICP one year from the previous one. The CF survey expectations are raw data, expressing monthly expectations for the average annual growth of consumer prices for the current year. The AR(12) forecasts are monthly forecasts, performed using estimated AR(12) processes for inflation, from which measures for the whole year are constructed: the line depicts the average annual inflation built using these auxiliary forecasts.

ANNEX A

Description of data

1. Introduction

Among possible surveys that could inform about the evolution of inflation expectations for the euro area, two have been retained: the European Commission consumer survey (EC) and the Consensus Forecasts (CF). Alternatives to these are the industry surveys produced by the European Commission itself, in which firms are asked to state their expected price decisions for the next three months, and the Survey of Professional Forecasters (SPF) of the ECB.

The present paper briefly describes the data and the procedures followed to derive expectations from them. The next section tackles the EC survey and section 3 does the same for the CF survey.

2. EC Consumer Survey

2.1 Description of the raw data

The EC consumer survey has already been the subject of a number of previous studies. The survey includes a number of questions on the personal economic situation of households, but also some on general economic conditions as perceived by households—inflation, unemployment and growth. Questions in the survey are of a qualitative nature, and must be quantified somehow. The quantification methodology followed in the present study is the same as the one in Forsells & Kenny (2002), itself based on Berk (1999). It is worth noting that the European Commission reports results from the survey (including inflation expectations) in the form of balances of opinions of qualitative answers, as will be explained later on. An analysis of the relationship between the adopted quantification and the balance of opinions is helpful in shedding light on the survey, and will be attempted at the end of the current section. But it is first important to describe the questions in the survey, as far as price expectations are concerned at least, and how these can be quantified.

The EC survey includes two questions which are relevant for the present analysis, a backward-looking one:

Q5 How do you think that consumer prices have developed over the last 12 months? They have …

1 risen a lot

2 risen moderately

3 risen slightly

4 stayed about the same

5 fallen

6 don't know,

and a forward-looking one:

Q6 By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months? They will …

1 increase more rapidly

2 increase at the same rate

3 increase at a slower rate

4 stay about the same

5 fall

6 don't know.

The survey is conducted each month in the EU member countries and refers to prices as generically perceived by households. Responses cannot thus be linked to any geographical zone, i.e. region or country. Data analysed corresponds to percentage of answers falling under each option for an aggregate of the euro area, for both questions Q5 and Q6. By construction, stated inflation expectations are taken as being for euro area inflation based on the HICP, although nothing in the survey allows to think that this is indeed the measure respondents have in mind. In the analysis that follows the time frame over which expectations are expressed will be assumed to be as year-on-year (i.e., 12 months ahead), although see Nielsen (2003) for alternative interpretations.

Both questions are charted in Figure 1 (Q5) and Figure 2 (Q6), in the form of euro area percentage of responses per category. Notable in the charts is the higher variability of percentages in Q5, compared to Q6; and the relatively large width of the "increase at the same rate" option in Q6. Another point worthy of note is the somewhat higher percentage of don't-know answers in Q6. Note that since this category of answers is ignored in the analysis, this remark has no bearing on the results that will be reported. Last but not least, both Q5 and Q6 show important swings in the latter part of the sample (2002 and 2003) which may have a bearing on the analysis.

2.2 Methodology

The methodology adopted to turn qualitative answers into explicit numerical inflation expectations is based on Berk (1999), itself derived from the approach by Carlson & Parkin (1975). The approach assumes responses to the survey are based on a specific distribution function. Households answering the survey use this distribution function to derive their responses, in the form of allocation to specific options in the questions based on the implicit quantiles of the distribution. Carlson and Parkin were able to solve, following a very simple approach, the conundrum of somehow fixing the thresholds used implicitly by the households in choosing specific answers. In their case, questions were of the type admitting three answers—of the sort “up”, “equal” or “down”. Their approach simply entailed transforming the percentage of answers for each option into z-scores of the assumed distribution, i.e. finding the threshold values of the distribution corresponding to the percentages. This was done by simply applying to the mentioned percentages the inverse of the distribution function. It is obviously necessary to derive implicit ranges for the "equal" option, since households cannot attach probabilities to a single point. This is done by simply solving a system of equations, in which the parameters needed to fully describe the distribution function and the ranges are obtained as the solution to a system of equations. Berk’s (1999) method is a simple adaptation of the procedure to the specific questions in the EC survey.

Take for instance question Q6, for which five categories are defined around two different benchmarks: zero inflation rate and current (perceived) inflation rate. Berk proposed to map responses to implicit percentiles of assumed distribution functions by allowing for implicit ranges for which respondents are not able to distinguish from specific points, i.e. zero or current inflation. The approach is graphically depicted in Figure 3, in which expected inflation π_{t+1}^{e} is represented by the mean of the distribution. Households are asked to choose options based on distance of expected inflation from (on the one hand) zero and (on the other hand) current perceived inflation π_t. Households do not make the comparison based on point benchmarks, but implicitly consider ranges around them, in the form of $[-\delta, +\delta]$ around zero and $[\pi_t - \mu, \pi_t + \mu]$ around current perceived inflation. The chart depicts the probability density function (PDF), although the approach uses the cumulated distribution function (CDF) to derive the z-scores.

Since Carlson and Parkin's method entails setting up a system of equations and solving it, it is important to map the questions in the survey into equations in a way reflecting their meaning. As we will see, there is no single way in which this mapping can be established for the EC survey. For instance, question Q5 above (about perceived inflation based on the past 12-month situation) can be understood as referring to current inflation compared to past expectations, or compared to historical averages, or maybe compared to recent past inflation. As to question Q6, it could refer to views about future inflation based on today’s inflation, or maybe based on perceptions of current inflation maybe different from its actual level. It is therefore important to properly define this mapping. For this, the two questions must be treated differently. Q5 asks about current prices compared to their level 12 months ago, letting respondents take as benchmark both zero inflation rate (prices staying the same) or some unspecified positive inflation rate (prices rising moderately). Question Q6 is even more explicit in setting two different benchmarks: zero inflation rate (prices staying about the same) and current inflation rate (prices increasing at the same rate). This leaves a rather large amount of possible choices to be made.

Note that the 'don't know' answer in both questions is not very informative: with few exceptions, the percentage of responses falling in this category is very stable. In the following, as has become standard in the literature, it is evenly distributed among the rest of answers.

One important last point is how to select the distribution function on which the quantification will be done. In the current study, the approach has been to adopt Normal distributions everywhere. First of all, due to a law-of-large-numbers consideration, i.e. the fact that the probabilities under analysis summarise views of a large group of households. Secondly, because results in the literature do not strongly support the use of other distributions, see for instance Nielsen (2003).

2.2.1 Question 5

Both questions are phrased in a way that deserves some thought. Question Q5 is not really clear about what the survey is really asking. It can be understood as asking about pre-conceived ideas about what inflation today should be compared to what it actually is or as asking about what are the perceptions of inflation today without a specific benchmark. Berk took the second meaning and derived a measure of perceived current inflation using the percentage of households choosing options 4 and 5 for the question and aggregating those that choose options 1 to 3, leading to generic categories 'prices increased', 'prices stayed the same' and 'prices fell'. In this sense, they just transform Q5 in order to enable direct application of the method by Carlson and Parkin. Berk then derived the mean and standard deviation using a formula reported below, and finally he estimated the bands around 0 inflation assuming that actual and perceived inflation had the same mean.

More formally, it is assumed that households answer based on an implicit probability for perceived inflation being 'negative', $prob\left(\pi^p - \delta \leq 0\right)$, for option 5 and being 'zero or lower', $prob\left(\pi^p + \delta \leq 0\right)$, for the cumulated responses for options 5 and 4. Options 1, 2 and 3 are bundled together in a ‘positive’ category, which is furthermore irrelevant since it amounts to the complement of the two previous categories.[1] This approach implies in practice ignoring any information included in the responses to options 1 and 2. An alternative will be explored below. (See also Nielsen (2003) for a similar point.)

Assuming a normal distribution for the answers, *z*-scores can be calculated for all the options: z_1 for option 1, z_2 for options 1 and 2 (cumulating the respective percentages), z_3 for options 1 to 3 and z_4 for options 1 to 4. Note that z_5, for options 1 to 5 is 1 by construction. Berk's approach would only use *z*-scores z_3 and z_4, discarding the others. Note that the *z*-scores are calculated for the complement of each option, i.e. using the inverse CDF of the Normal distribution for 1 minus the cumulated percentages for each option. Berk calculates the mean and standard deviation using:

$$m = -\delta \frac{z_3 + z_4}{z_3 - z_4},$$

$$s = \frac{2\delta}{z_3 - z_4},$$

and setting δ such that the calculated mean of expectations coincides with the sample one for inflation. Note that the calculations are done for each period, leading to mean and standard deviation as time series.

The resulting mean is graphed in Figure 4, together with year-on-year HICP inflation rate and bands in the form of 2 standard deviations. It is evident that a problem with the method appears towards the end of 2002 and in 2003. The problem with this approach is that it depends on the width of the bin at the

---

[1] Remember that the “don’t know” category is ignored and the other options are re-scaled to add up to 100.

extreme, i.e. it depends on whether a relatively sizeable proportion of respondents has chosen option 5. If this is not the case, and only a handful of households chose it in some periods, mean and standard deviation will be imprecisely estimated. In all likelihood, this is what happens in 2003: the number of households reporting an expected fall in prices decrease significantly and this leads to a very narrow 'fall' bin. In this case z_3 and z_4 are very similar and the expression for m becomes unboundedly large.

There are a number of alternatives to solving this problem. One is to acknowledge that the problem stems probably from a problematic period for the survey, and simply smooth the spike by interpolation. Other option is to somehow take instead the other, much larger category bins. For instance, assuming that households split the space evenly, i.e. that the distance from zero in option 3 is twice the distance from zero in option 4, and the distance from zero in option 2 is three times that. Using this assumption, mean and standard deviation can be calculated using three different sets of formulas: Berk's original one, another set taking z_2 and z_3, and finally a third one taking z_1 and z_2.

The first alternative would lead to

$$m = \delta \frac{z_2 - 2z_3}{z_2 - z_3},$$

$$s = \frac{\delta}{z_2 - z_3}.$$

The second to

$$m = \delta \frac{2z_1 - 3z_2}{z_1 - z_2},$$

$$s = \frac{\delta}{z_1 - z_2}.$$

The first alternative is charted in Figure 5, the second in Figure 6. These alternatives avoid the previous problem and lead to similar estimates as previously for the periods previous to 2003, at least for the mean. Further sophistication can be achieved, but it is not clear whether it would lead to very different results.

A third option, adopted by Nielsen (2003) and worth pursuing, is to make full use of the five categories and adopt an approach similar to the one explained in the next sub-section, as applied to question Q6. This method delivers smooth estimates of perceived inflation, contrary to Berk’s original approach, which are actually close to perceived inflation derived using actual past inflation. The problem with this approach (if it can be described as such at all) is that questions Q5 and Q6, the latter analysed in the next section, are not identical: a mapping reflecting answers for Q6 may be inappropriate as a vehicle to quantify Q5.

Note that the resulting smoothing is not complete: there is in all measures some degree of hump shape for perceived inflation in 2002/2003. Just looking at the answers for Q5 overt his period, Figure 1, it can be easily seen that there is indeed at the time a compression of beliefs towards perceived much higher inflation. This could be a result of the introduction of the euro and the well-publicised belief among the public that this led to generalised hikes in prices. As will be seen, this episode is also visible in the case of professional forecasters, although in their case to a much lesser extent. Unfortunately, the sample used in this study precludes any serious analysis of the euro cash changeover, and should be considered out of the scope of the paper. The point, tackled in Ehrmann (2006) and, from a policy perspective, in the September 2006 issue of the ECB Monthly Bulletin, is certainly worth further analysis.

2.2.2 Question 6

Question 6 is phrased in such a way that the benchmark value households have in mind is current inflation, which they compare with their expectations. Current inflation can be taken to be last released inflation, i.e. the publicly available year-on-year inflation datum in the month previous to the conduct of the survey. An alternative is to take the perceived inflation coming from question 5, as explained above, to use as the benchmark. The two measures have been taken to derive a quantitative measure for inflation expectations, as charted on Figure 7 and Figure 8. The charts depict expected inflation taking, correspondingly, the first and second approaches mentioned previously (shifted in time to accommodate the timing of the forecasts), together with actual HICP inflation and 2-standard-deviation bands.

Formulas used to derive the two measures are explained in the appendix of Berk (1999). The approach is an extension of Carlson and Parkin adapted to the specific question Q6 of the EC consumer survey. Where Carlson and Parking were expecting questions answerable with three categories, Berk extends the analysis to five categories and two different thresholds. The approach has the advantage that the mean expected inflation can be derived using only cross-section information, i.e. expectations must not be re-scaled using time series information.

The approach involves calculating the *z*-scores of the cumulated options, as before. The percentages reported for each release of the survey are cumulated (so that for option *i* the retained percentage is that of respondents having answered any of options 1 up to *i*). For each percentage, the inverse Normal (normalised to zero mean, unit variance) is applied to one less the given percentage, i.e. $z_i = \mathrm{CDF}^{-1}_{\mathrm{Normal}}\left(100 - \%\ of\ \ cases\ where\ response \le i\right)$. Thus, the *z*-scores z_1 to z_5 are obtained, where again z_5 can be dropped because it is unity by construction. The estimates of *m*, the mean expected inflation, *s*, its standard error, δ, the threshold to consider expected inflation to differ from zero, and μ, the threshold to consider expected inflation different from perceived inflation, can be derived from the following system of equations

$$\frac{\pi^p + \mu - m}{s} = z_1$$

$$\frac{\pi^p - \mu - m}{s} = z_2$$

$$\frac{0 - \delta - m}{s} = z_3$$

$$\frac{0 - \delta - m}{s} = z_4$$

Where π^p can be either taken from the quantification of question Q5 or from past observed inflation. The system is solved by:[2]

$$m = -\pi^p \frac{z_3 + z_4}{z_1 + z_2 - z_3 - z_4}$$

$$s = \frac{2\pi^p}{z_1 + z_2 - z_3 - z_4}$$

$$\mu = -\pi^p \frac{z_1 - z_2}{z_1 + z_2 - z_3 - z_4}$$

$$\delta = \frac{\pi^p (z_3 - z_4)}{z_1 + z_2 - z_3 - z_4}$$

The relative smoothness across time of the answers leads to smoothly varying inflation expectations. This is perceived in Figure 7 in the lagging behaviour of expectations compared to realisations, a mechanical effect of the fact that expectations in the chart were calculated based on last month's inflation. In Figure 8 the smoothness is enhanced by the fact that perceived inflation as derived from Q5, used to calculate this version of the measure, is itself very smooth.

2.3 Balance of opinions for the two questions

Eurostat does not report percentage of answers but calculates a statistic termed 'balance of opinions', as the weighted average of percentages in each answer. Calling p_i the percentage of responses corresponding to option *i*, the balance of opinions for questions Q5 and Q6 is calculated as p_1+0.5 p_2-0.5 p_4-p_5, thus giving a weight of 1 to option 1, 0.5 to option 2, 0 to option 3, -0.5 to option 4 and –1 to option 5. One interesting feature of this index is how closely it tracks at times the expectation series derived previously, see Figure 9. The single notable exception is the latter part of the sample, i.e. 2002 and 2003, for which important discrepancies appear between the two. For question Q5, for instance, the balance of opinion is much higher in this period than the corresponding series for perceived inflation extracted previously. The contrary happens for Q6, for which the latter part of the sample is systematically below its quantification.

This fact may raise some doubts on the usefulness of Berk's method, when simpler mechanisms seem to lead to similar constructs. In fact, it will be argued, there could be circumstances in which Eurostat's index may provide noisy signals. To gain an intuitive understanding of this point, the balance of opinions could be seen as a random variable whose mean is being calculated. This assumed random variable would only take five different values (-1, -0.5, 0, 0.5 and 1) and the probability of it reaching any of the values would be given by the probabilities in the survey. The calculation of the mean can then be decomposed as in (1), in which the support for variable *z* has been split in five parts, each corresponding to regions where the variable takes each of the assumed five values. Note that in the expression the value taken by the

[2] Replication files cannot unfortunately be released because of the confidential nature of the data. As an illustration of the simplicity of the technique, nevertheless, Matlab code in Annex B documents calculations with a dummy survey. The code uses no data (as the survey data is simulated), so it can be pasted directly into Matlab and run.

integrals is known—it corresponds to the declared probabilities—and the integration limits do not therefore need to be known. This leads to a great simplification in the calculations.

$$I = \int_{-\infty}^{+\infty} z\,\phi(z)\,dz = \\ = 1\int_{-\infty}^{z_1}\phi(z)\,dz + 0.5\int_{z_1}^{z_2}\phi(z)\,dz + 0\int_{z_2}^{z_3}\phi(z)\,dz - 0.5\int_{z_3}^{z_4}\phi(z)\,dz - 1\int_{z_4}^{+\infty}\phi(z)\,dz \qquad (1)$$

In sharp contrast, perceived inflation and inflation expectations have been extracted assuming a standard Normal distribution, implying a continuous random variable. Therefore, the space cannot be split and the integration limits matter. It is nevertheless intuitive that there is a case for which the two calculations can lead to a similar profile: when categories lie symmetrically with respect to the mean of perceived or expected inflation. This can happen if, for instance, perceived (expected) inflation were to lie exactly in the middle of zero and past (current) inflation. Another possibility would be of stable but high inflation: in this case, perceived (expected) inflation would coincide with past (current) inflation, and the probability attached to attaining zero inflation would be zero. On the other hand, in a situation like that recorded in 2002/2003, in which perceived inflation is much higher than past inflation, the weighing implied in (1) is such that the shift is very much exaggerated. This seems to be the case at the end of the sample, as seen in Chart 9, when the spike in perceived inflation and the corresponding drop in future expected inflation are excessive compared to our measure. In conclusion, there are circumstances in which Eurostat's index could give misleading indications.

3. Consensus Forecasts

The Consensus Forecasts are collected from professional forecasters on a monthly basis, from January 1990. Forecasts are for current- and next-year inflation (i.e., only annual inflation is forecast), and refer to the country of the forecaster before 1999, after which a question specifically for the euro area was included in the survey. Series used in the analysis (covering only the euro area in aggregate) are for the area as a whole for the period starting in January 1999, and correspond to the weighted average of the corresponding country information for the periods before that date. The last survey included was for September 2003.[3]

Since forecasts are for full years, the treatment of the data must be based on the month in which each wave of the survey was launched. For instance, a survey conducted in January (for which it is assumed that the January HICP had not yet been released) must be understood as implying a 12-month-ahead forecast for the current year, only known one year later, and 24-month-ahead forecast for the next year. Since there is no trivial way to interpolate forecasts to lower frequencies, there is in practice no alternative to this approach of treating surveys separately according to the month they were carried out. An overview of the CF forecasts can be gained in Figure 10 (current year) and Figure 11 (next year) against actual HICP annual inflation, distinguishing by the month in which the survey was carried out.

[3] Data was kindly lent by Magnus Forsells.

One important element in the graph is the evidence of higher expected inflation in 2001, both for current- and next-year forecasts, in line with the EC survey although less dramatically. It is interesting to observe that forecasts for the year 2001 evolve in a very significant way: from an average forecast lower than actual inflation in the January 2001 survey, forecasts slowly creep upwards until they cross actual inflation as of the June 2001 survey, staying afterwards well above final realised inflation.

Figure 1. EC consumer survey, question Q5: current prices, compared to 12 months ago, have…

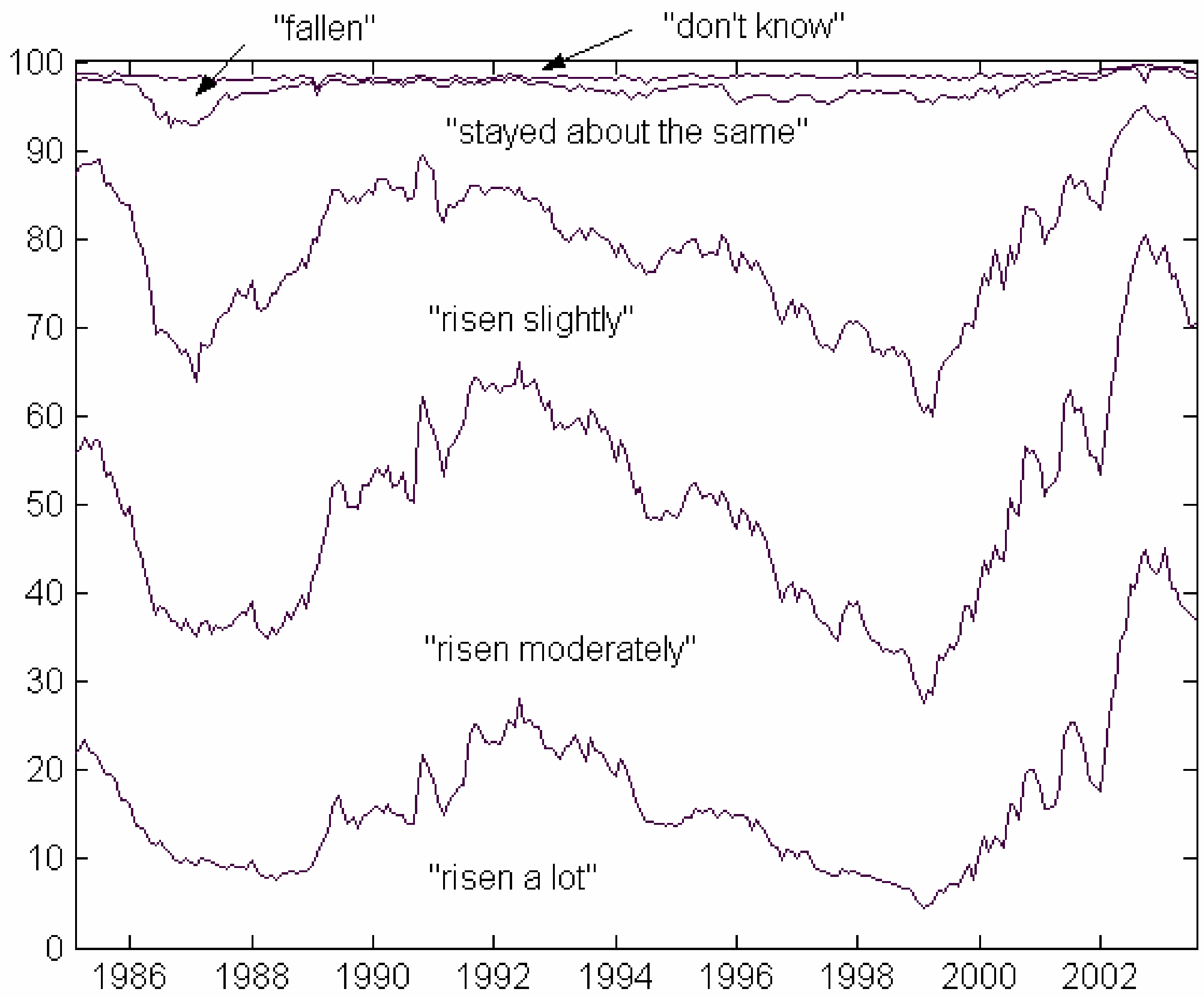


Figure 2. EC consumer survey, question Q6, prices over the next 12 months will…

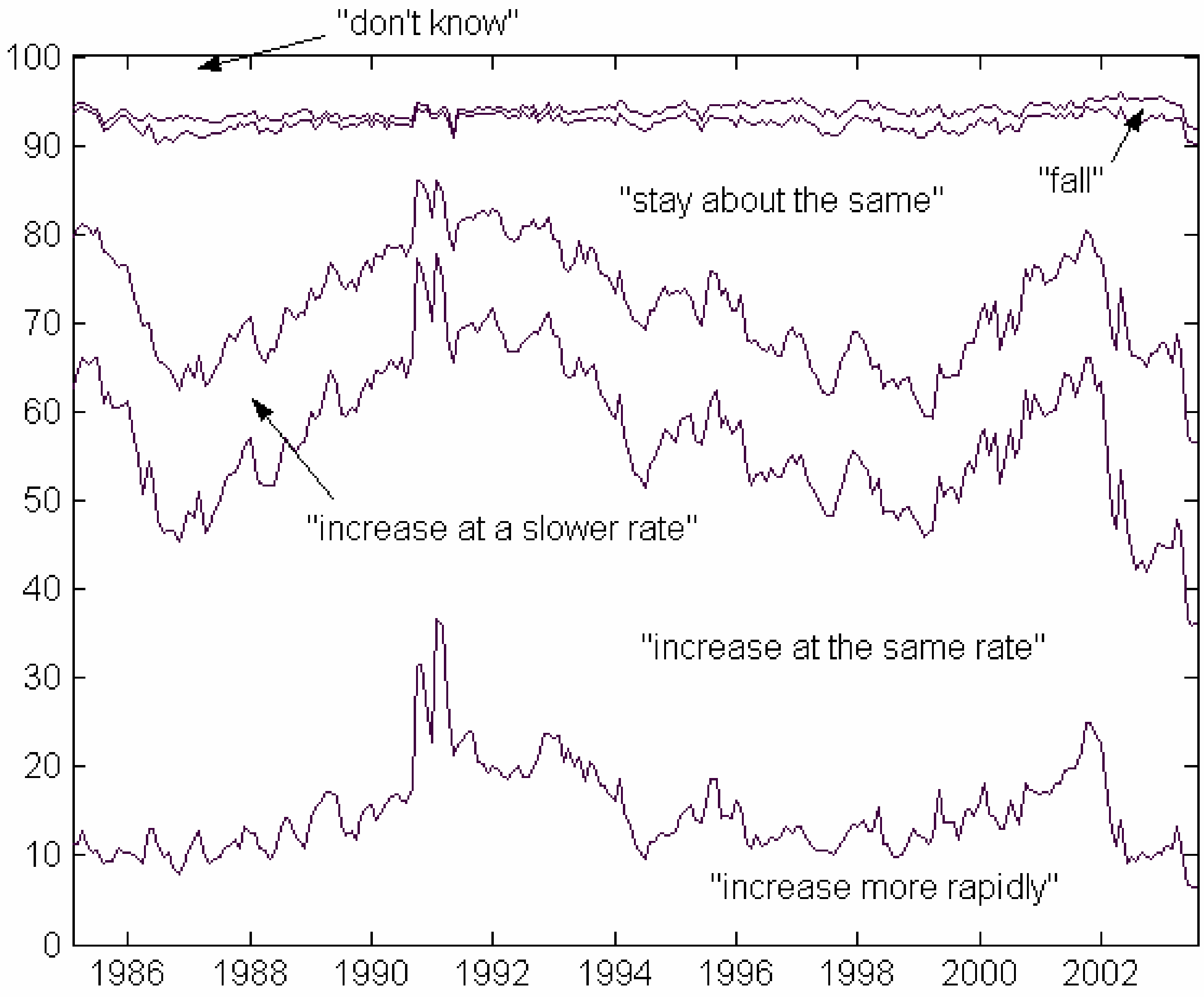


Figure 3. Berk's Quantification Framewrok

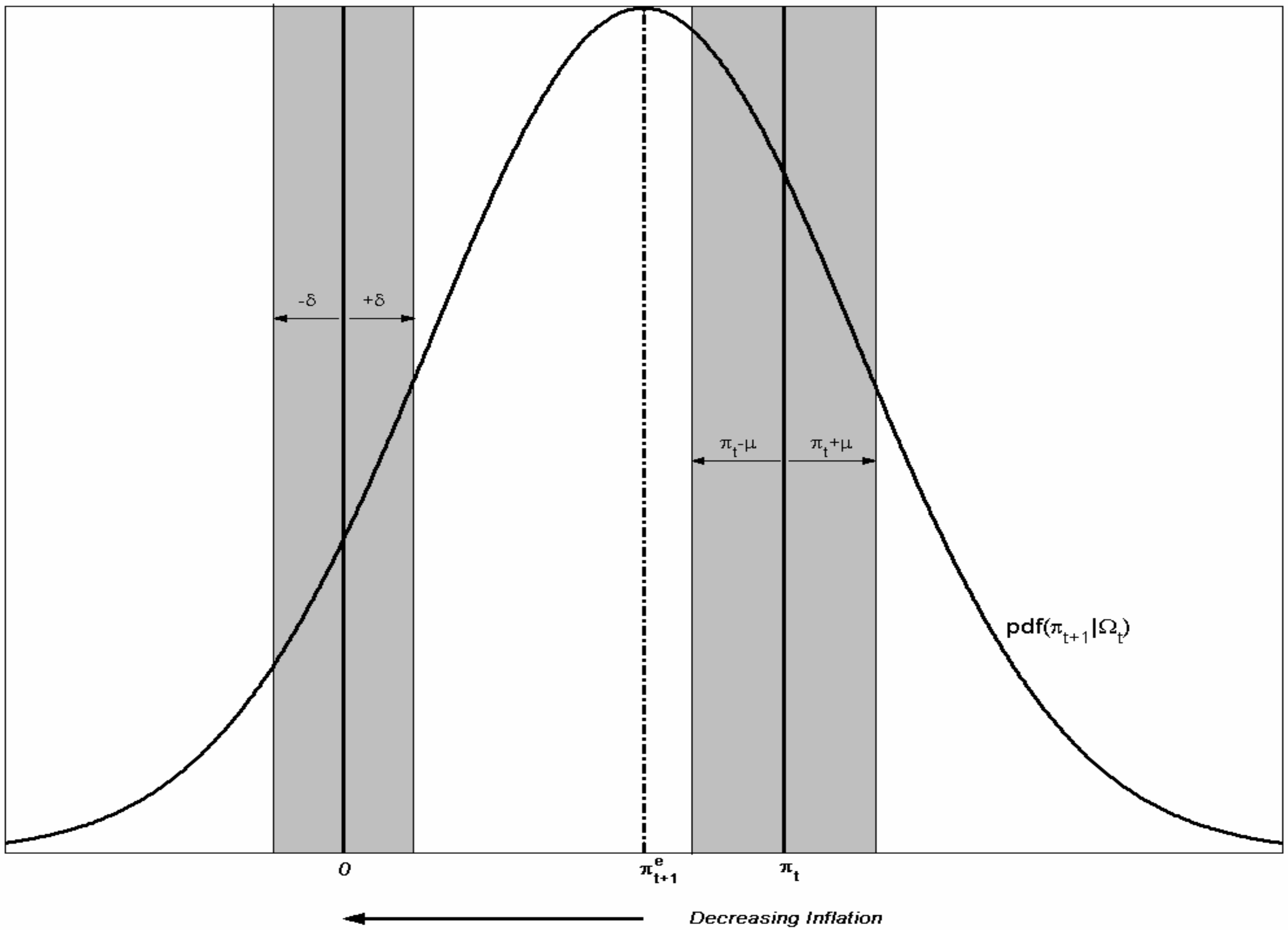


Figure 4

Perceived Inflation, Original Berk's Method

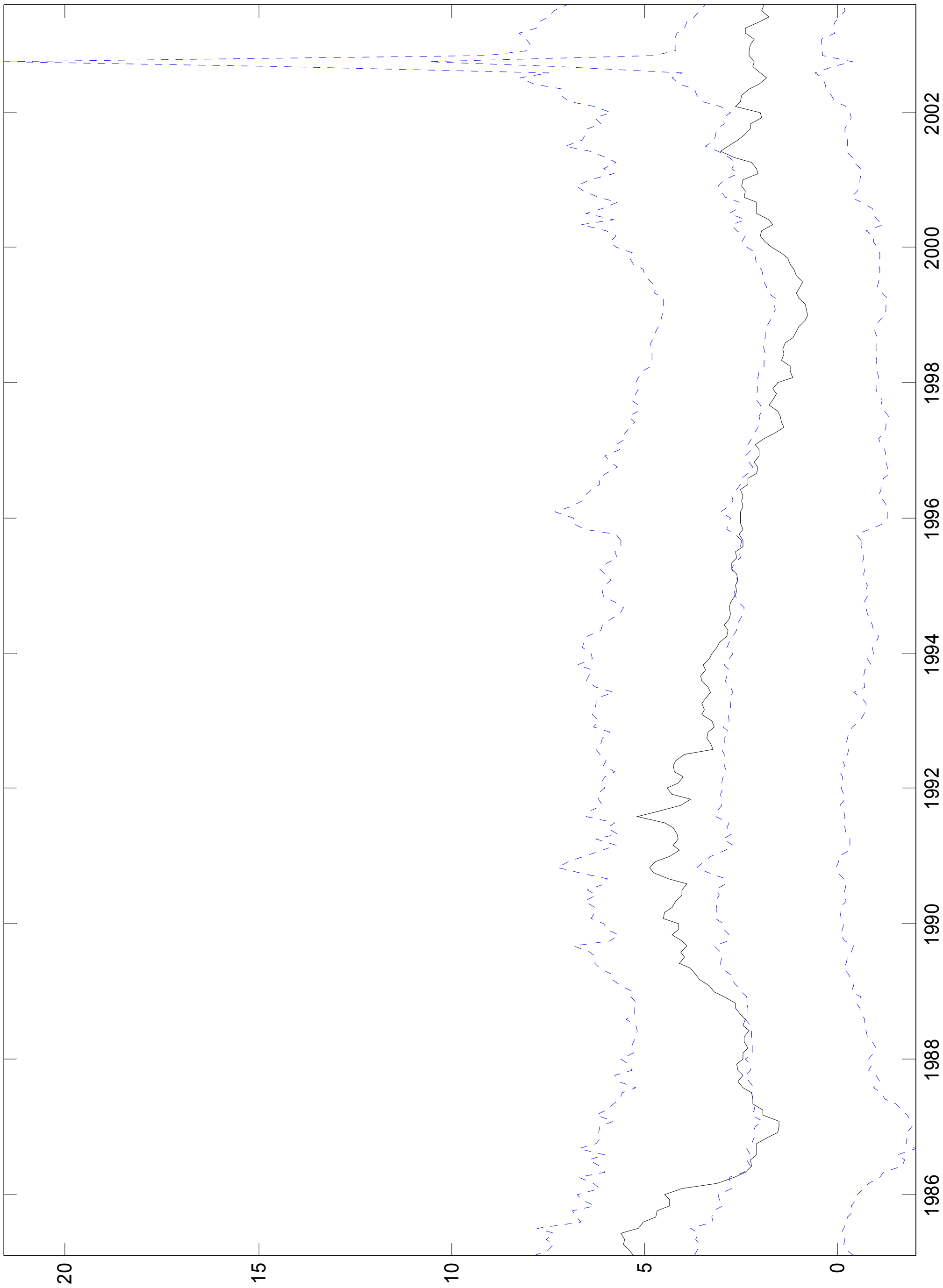


Figure 5

Perceived Inflation, Modified Berk's Method (I)

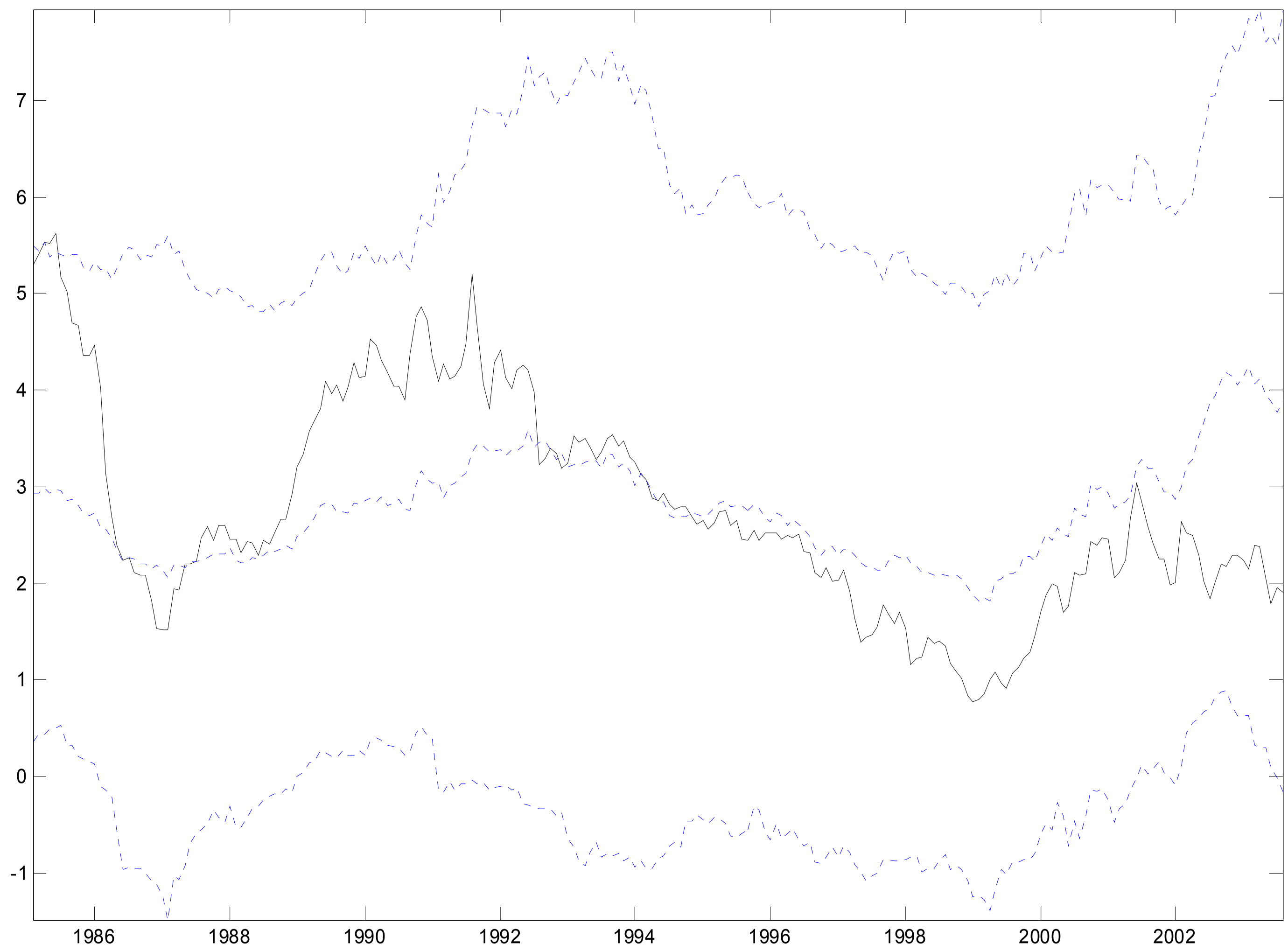


Figure 6
Perceived Inflation, Modified Berk's Method (II)

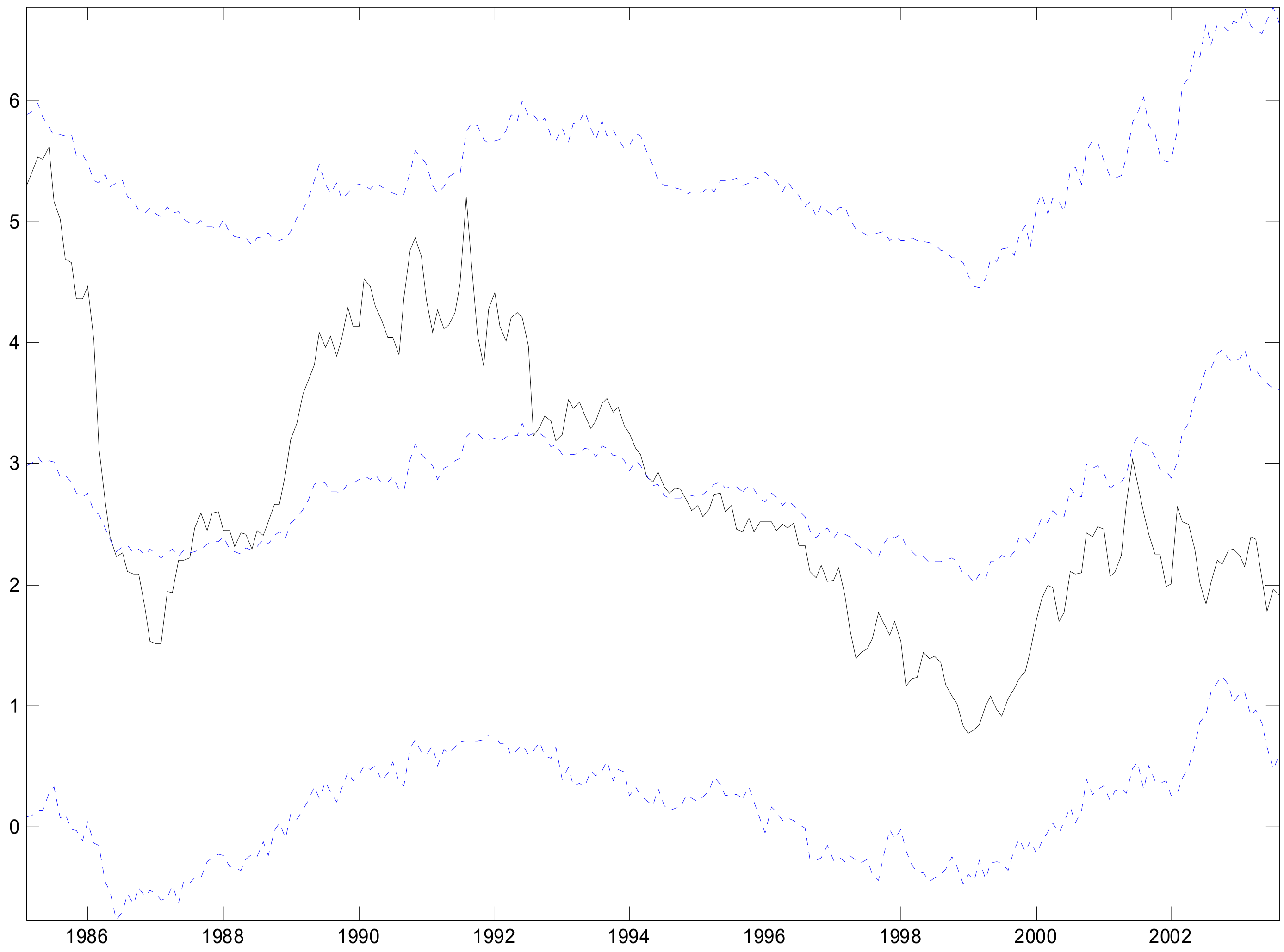


Figure 7
Expected Inflation (I)

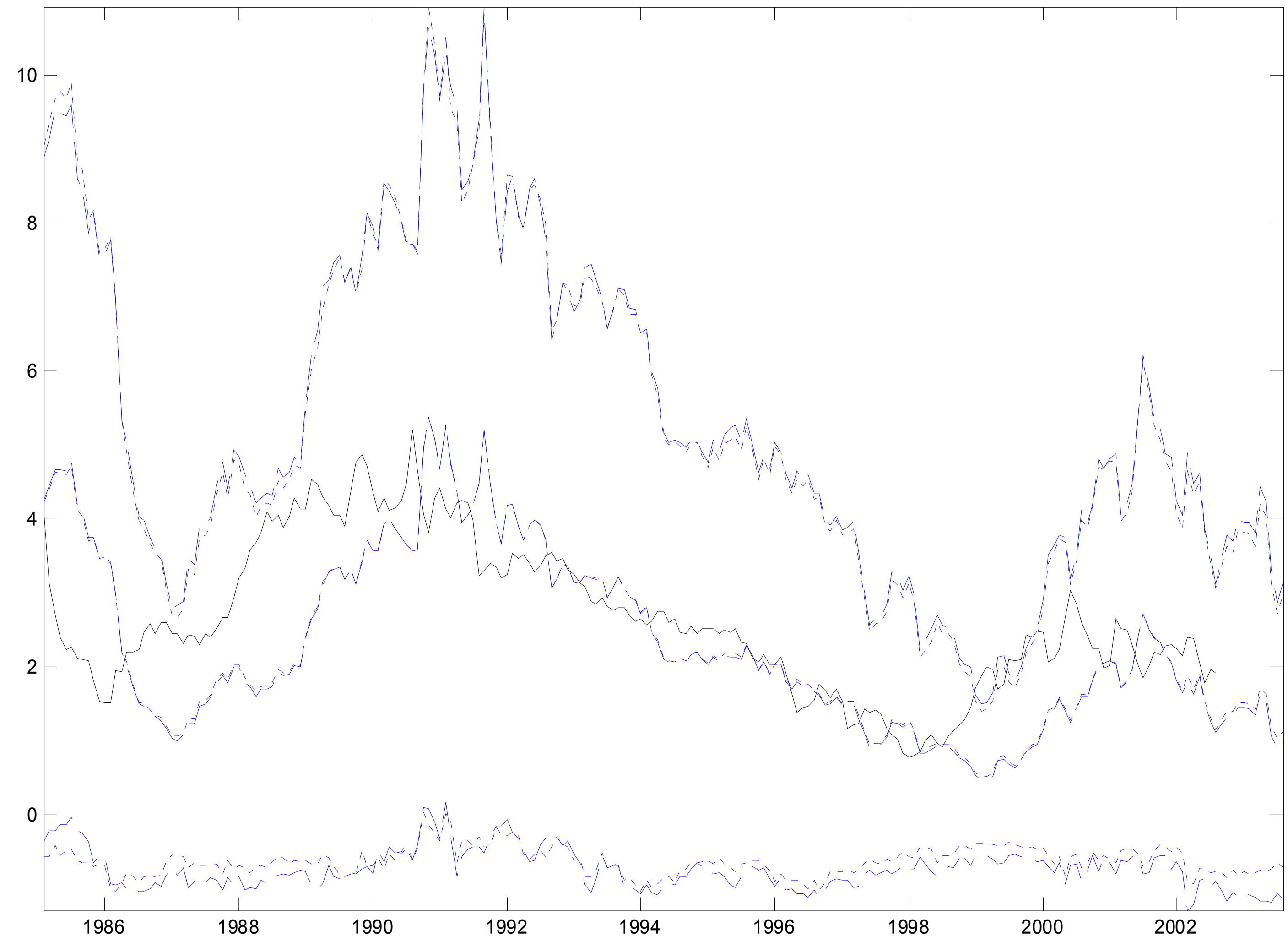


Figure 8

Expected Inflation (II)

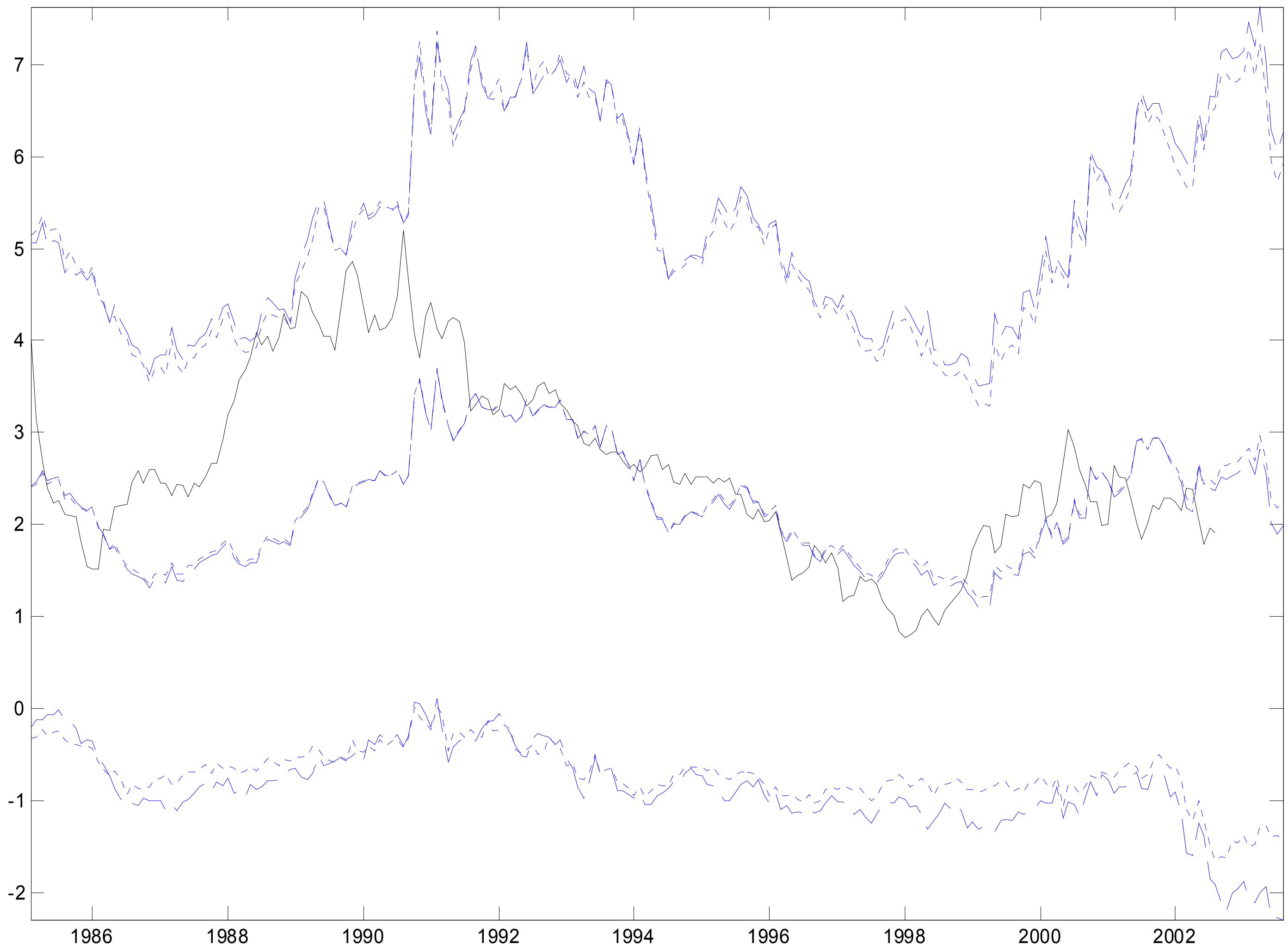
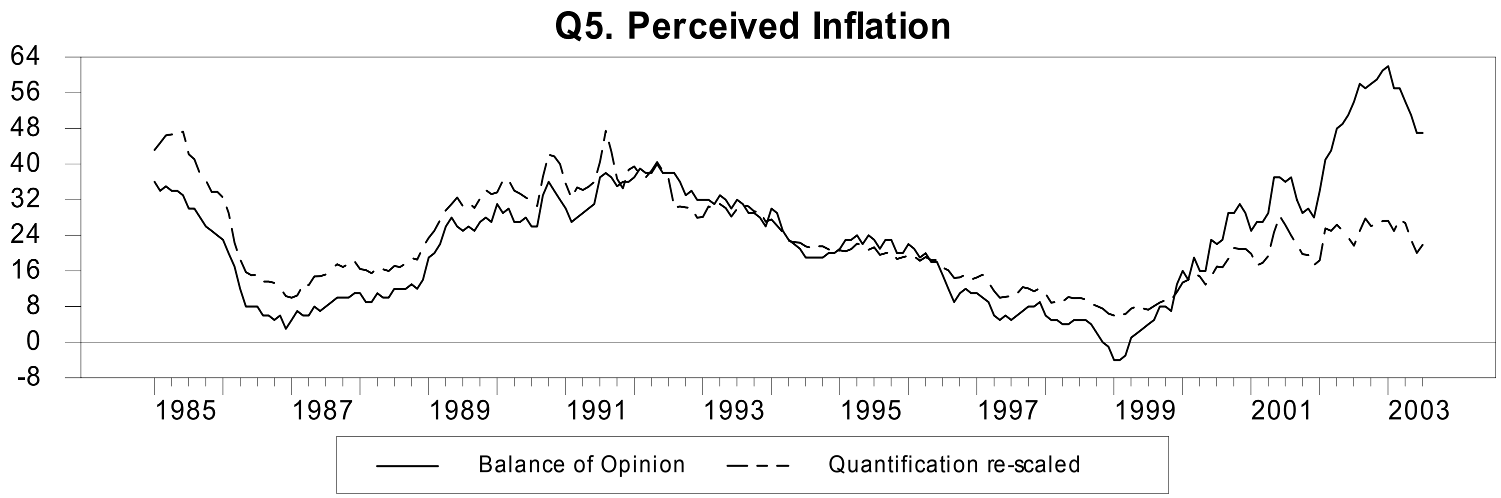


Figure 9. Balance of Opinions vs. Quantified Surveys



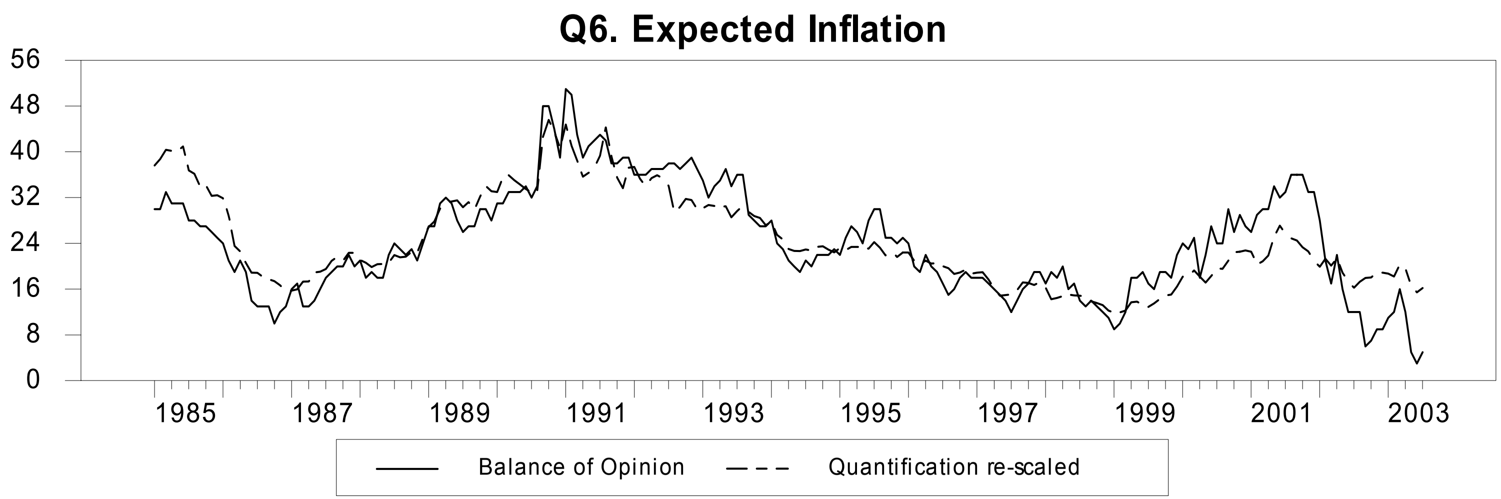


Figure 10

Consensus Forecasts

Inflation Expectations, Current Year

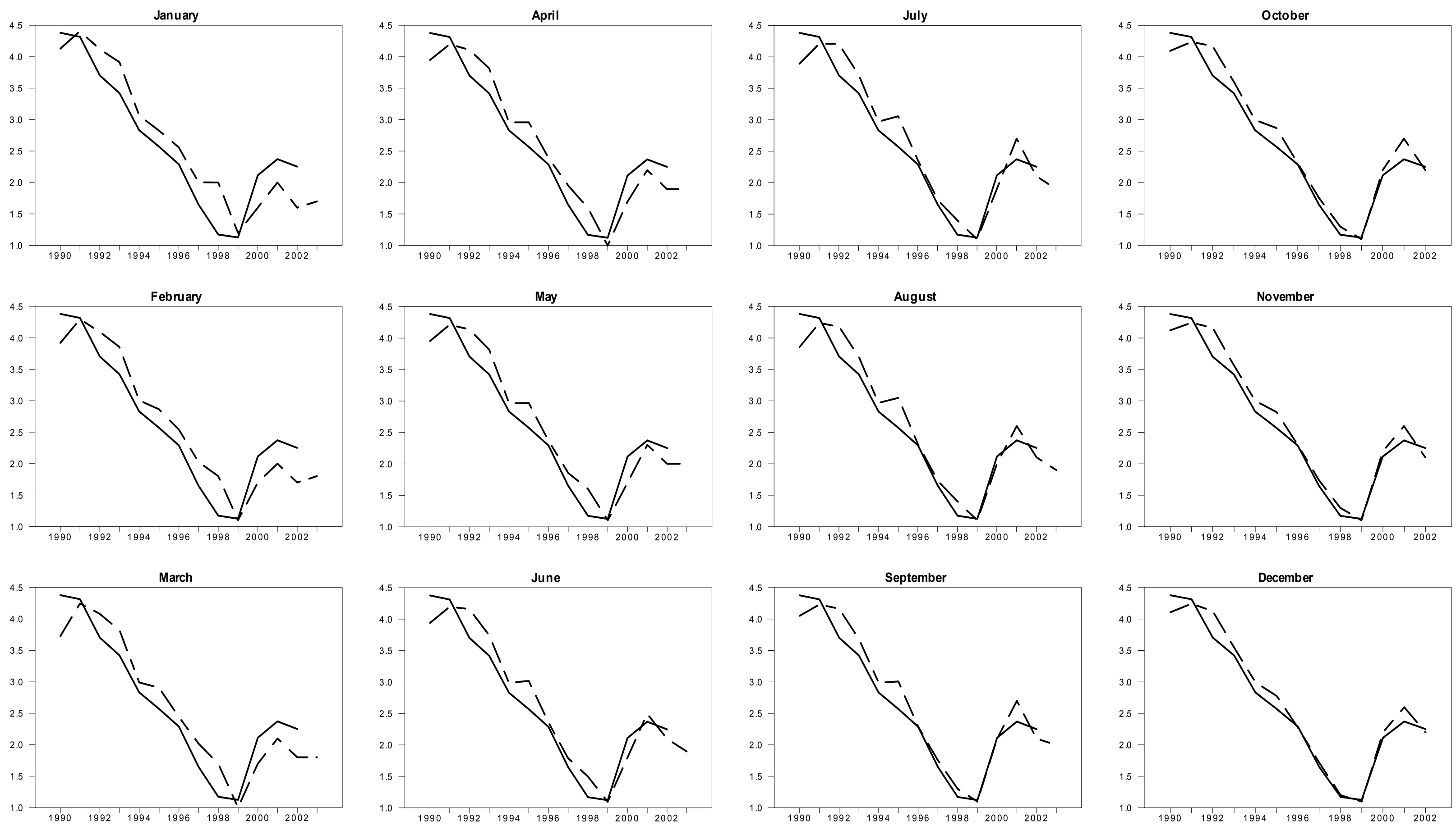
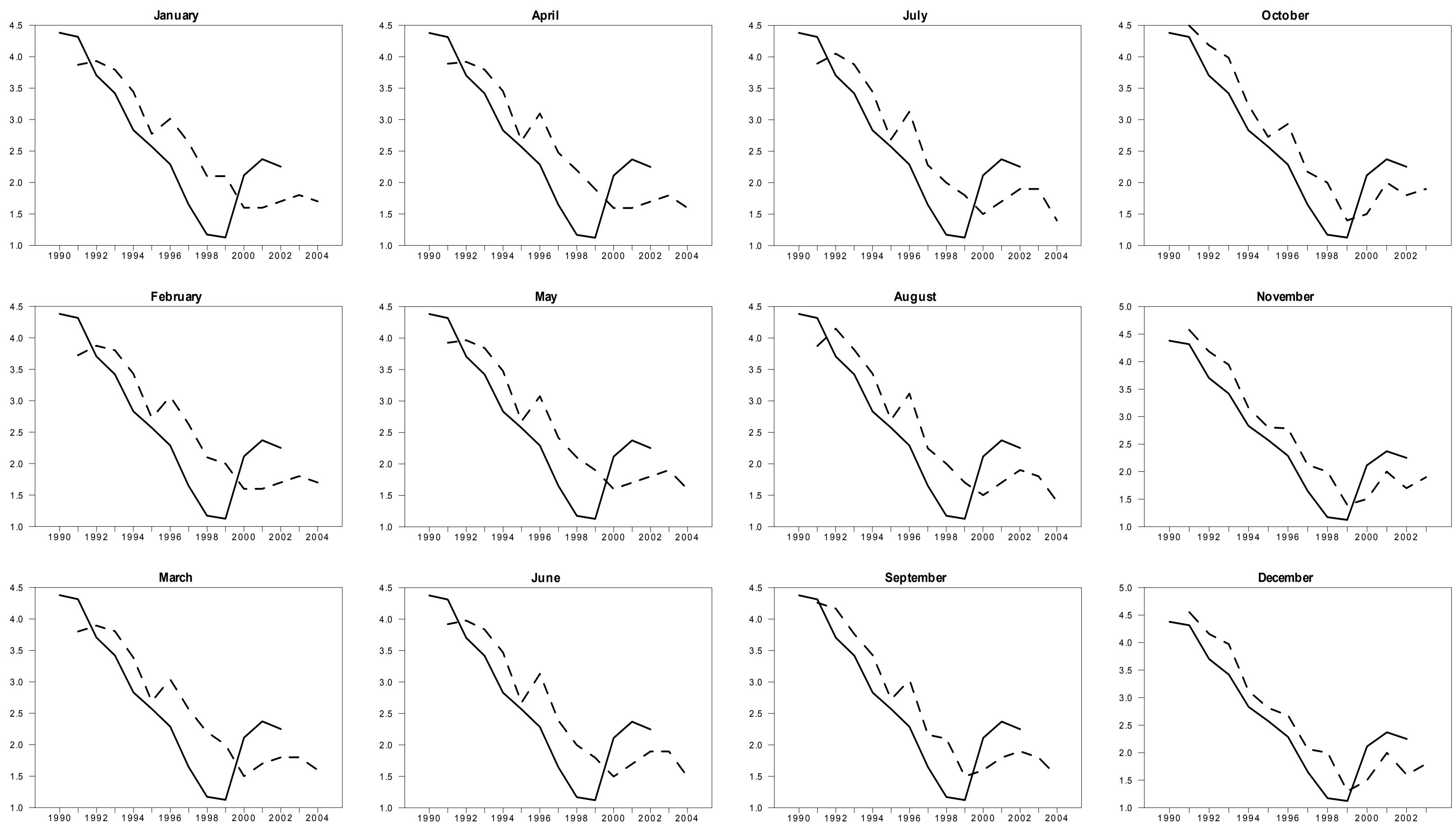


Figure 11

Consensus Forecasts

Inflation Expectations, Next Year



ANNEX B

Dummy Matlab code

1. Matlab code with dummy survey

```
% CreateSurvey.m
% Creates artificial draw for survey and checks Carlson-Parking and Berk
clear all
close all
clc

% To be able to replicate draws
randn('state',0)

% Treatment for question Q6 -- Expected inflation

% This is the observed perceived inflation
PerceivedInflation=2.5;

% NB: note that a non-observed, stochastic perceived inflation
% would make for less precise estimates of mean expected
% inflation, but would not change the asymptoptics.

% These are the unobserved population parameters
MeanExpectedInflation=2;
StdErrorExpectedInflation=1.25;

% This is the size of a launch of the survey
Observations=50000;

% In the next few lines,we generate the actual inflation expectations.
% They are assumed to be drawn from a single Normal distribution. By
% commenting and uncommenting lines below, a t distribution can be selected
% instead.

% Standard case: normal distribution
ExpectedInflation=randn(Observations,1)* ...
    StdErrorExpectedInflation+MeanExpectedInflation;

% % Non-standard case: t distribution (fat tails)
% ExpectedInflation=trnd(2,Observations,1)* ...
%     StdErrorExpectedInflation+MeanExpectedInflation;

% Mapping of draws into responses

% First, set thresholds for answers (common to all respondents)
DELTA=0.2;     % distance to zero inflation (shared by all respondents)
MU=0.25; % distance to perceived inflation (shared by all respondents)

% Then, set the (unreported) individual responses according to codes in survey
IndividualResponses=1+ ...
    (ExpectedInflation<PerceivedInflation+MU)+ ...
    (ExpectedInflation<PerceivedInflation-MU)+ ...
    (ExpectedInflation<DELTA)+ ...
    (ExpectedInflation<-DELTA);

% Finally, set the reported percentage of responses per category (cumulated)
Percentage=100*cumsum([ ...
    sum(IndividualResponses==1) ...
    sum(IndividualResponses==2) ...
```

```
    sum(IndividualResponses==3) ...
    sum(IndividualResponses==4) ...
    sum(IndividualResponses==5) ...
    ])./Observations;

% Now, calculate the z-scores
% Beware, Statistics toolbox needed!
z1=norminv(1-Percentage(1)./100);
z2=norminv(1-Percentage(2)./100);
z3=norminv(1-Percentage(3)./100);
z4=norminv(1-Percentage(4)./100);

%
% Carlson-Parkin (mean expected inflation assumed known)
%
% NB: note that Carlson-Parkin, contrary to Berk, use time-dimension
% information to re-scale m. Since there is no time dimension in this
% example, we will assume the mean expected inflation is known.
m=-(z3+z4)./(z3-z4);
DELTAhat=MeanExpectedInflation./mean(m);
m=m*DELTAhat;
s=2*DELTAhat./(z3-z4);

fprintf('*\n* Reporting Carlson-Parkin Results\n*\n');
fprintf('* NB: Actual and estimated means equal by construction\n');
fprintf('Distance to Zero Inflation\nActual: %6f; Estimated: %6f\n', ...
    DELTA,DELTAhat);
fprintf('Mean of Perceived Inflation\nActual: %6f; Estimated: %6f\n', ...
    MeanExpectedInflation,m);
fprintf('Std. Error of Perceived Inflation\nActual: %6f; Estimated: %6f\n', ...
    StdErrorExpectedInflation,s);

%
% Berk
%
m=-PerceivedInflation*(z3+z4)./(z1+z2-z3-z4);
MUhat=PerceivedInflation*(z1-z2)./(z1+z2-z3-z4);
DELTAhat=PerceivedInflation*(z3-z4)./(z1+z2-z3-z4);
s=2*PerceivedInflation./(z1+z2-z3-z4);

fprintf('*\n* Reporting Berk Results\n*\n');
fprintf('Distance to Zero Inflation\nActual: %6f; Estimated: %6f\n', ...
    DELTA,DELTAhat);
fprintf('Distance to Perceived Inflation\nActual: %6f; Estimated: %6f\n', ...
    MU,MUhat);
fprintf('Mean of Perceived Inflation\nActual: %6f; Estimated: %6f\n', ...
    MeanExpectedInflation,m);
fprintf('Std. Error of Perceived Inflation\nActual: %6f; Estimated: %6f\n', ...
    StdErrorExpectedInflation,s);
```

2. Resulting output file

```
*
* Reporting Carlson-Parkin Results
*
* NB: Actual and estimated means equal by construction
Distance to Zero Inflation
Actual: 0.200000; Estimated: 0.193486
Mean of Perceived Inflation
Actual: 2.000000; Estimated: 2.000000
Std. Error of Perceived Inflation
Actual: 1.250000; Estimated: 1.252099
*
* Reporting Berk Results
*
Distance to Zero Inflation
```

```
Actual: 0.200000; Estimated: 0.194565
Distance to Perceived Inflation
Actual: 0.250000; Estimated: 0.252720
Mean of Perceived Inflation
Actual: 2.000000; Estimated: 2.011156
Std. Error of Perceived Inflation
Actual: 1.250000; Estimated: 1.259083
```

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.int)

691 "The yield curve as a predictor and emerging economies" by A. Mehl, November 2006.

692 "Bayesian inference in cointegrated VAR models: with applications to the demand for euro area M3" by A. Warne, November 2006.

693 "Evaluating China's integration in world trade with a gravity model based benchmark" by M. Bussière and B. Schnatz, November 2006.

694 "Optimal currency shares in international reserves: the impact of the euro and the prospects for the dollar" by E. Papaioannou, R. Portes and G. Siourounis, November 2006.

695 "Geography or skills: What explains Fed watchers' forecast accuracy of US monetary policy?" by H. Berger, M. Ehrmann and M. Fratzscher, November 2006.

696 "What is global excess liquidity, and does it matter?" by R. Rüffer and L. Stracca, November 2006.

697 "How wages change: micro evidence from the International Wage Flexibility Project" by W. T. Dickens, L. Götte, E. L. Groshen, S. Holden, J. Messina, M. E. Schweitzer, J. Turunen, and M. E. Ward, November 2006.

698 "Optimal monetary policy rules with labor market frictions" by E. Faia, November 2006.

699 "The behaviour of producer prices: some evidence from the French PPI micro data" by E. Gautier, December 2006.

700 "Forecasting using a large number of predictors: Is Bayesian regression a valid alternative to principal components?" by C. De Mol, D. Giannone and L. Reichlin, December 2006.

701 "Is there a single frontier in a single European banking market?" by J. W. B. Bos and H. Schmiedel, December 2006.

702 "Comparing financial systems: a structural analysis" by S. Champonnois, December 2006.

703 "Comovements in volatility in the euro money market" by N. Cassola and C. Morana, December 2006.

704 "Are money and consumption additively separable in the euro area? A non-parametric approach" by B. E. Jones and L. Stracca, December 2006.

705 "What does a technology shock do? A VAR analysis with model-based sign restrictions" by L. Dedola and S. Neri, December 2006.

706 "What drives investors' behaviour in different FX market segments? A VAR-based return decomposition analysis" by O. Castrén, C. Osbat and M. Sydow, December 2006.

707 "Ramsey monetary policy with labour market frictions" by E. Faia, January 2007.

708 "Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting" by I. Vansteenkiste, January 2007.

709 "Quantifying and sustaining welfare gains from monetary commitment" by P. Levine, P. McAdam and J. Pearlman, January 2007.

710 "Pricing of settlement link services and mergers of central securities depositories" by J. Tapking, January 2007.

711 "What "hides" behind sovereign debt ratings?" by A. Afonso, P. Gomes and P. Rother, January 2007.

712 "Opening the black box: structural factor models with large cross-sections" by M. Forni, D. Giannone, M. Lippi and L. Reichlin, January 2007.

713 "Balance of payment crises in emerging markets: How early were the "early" warning signals?" by M. Bussière, January 2007.

714 "The dynamics of bank spreads and financial structure" by R. Gropp, C. Kok Sørensen and J.-D. Lichtenberger, January 2007.

715 "Emerging Asia's growth and integration: How autonomous are business cycles?" by R. Rüffer, M. Sánchez and J.-G. Shen, January 2007.

716 "Adjusting to the euro" by G. Fagan and V. Gaspar, January 2007.

717 "Discretion rather than rules? When is discretionary policy-making better than the timeless perspective?" by S. Sauer, January 2007.

718 "Drift and breaks in labor productivity" by L. Benati, January 2007.

719 "US imbalances: the role of technology and policy" by R. Bems, L. Dedola and F. Smets, January 2007.

720 "Real price wage rigidities in a model with matching frictions" by K. Kuester, January 2007.

721 "Are survey-based inflation expectations in the euro area informative?" by R. Mestre, February 2007.

ISSN 1561081-0

9 771561 081005