

EUROPEAN CENTRAL BANK  
**WORKING PAPER SERIES**



**WORKING PAPER NO. 33**

**THE INFORMATION  
CONTENT OF M3 FOR  
FUTURE INFLATION**

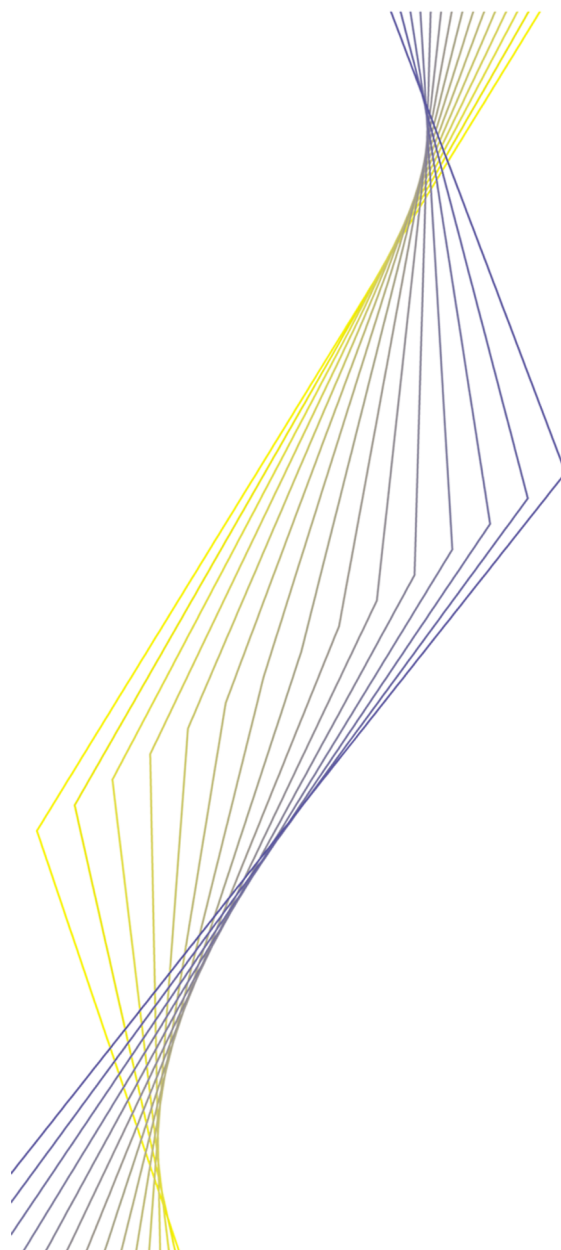
**BY CARMINE TRECROCI AND  
JUAN LUIS VEGA**

**October 2000**





**EUROPEAN CENTRAL BANK  
WORKING PAPER SERIES**



**WORKING PAPER NO. 33**

**THE INFORMATION  
CONTENT OF M3 FOR  
FUTURE INFLATION\***

**BY CARMINE TRECROCI AND  
JUAN LUIS VEGA**

**October 2000**

\* This research was carried out while the first author was at the ECB as part of the Research Graduate Programme run by the Directorate General Research. We are particularly indebted to S. Gerlach, L.E.O. Svensson, M. Kuismanen, M. Scharnagl and an anonymous referee for insightful discussion of our results. We also wish to thank I. Angeloni, J. Darby, V. Gaspar, J. Malley, A. Muscatelli, R. MacDonald, A.H. Meltzer, R. Sandilands, F. Smets, F. Spinelli, J. Stock, S. Wren-Lewis and U. Woitek for comments. Finally, this work has also benefited from suggestions by members of the ECB Working Group on Econometric Modelling as well as seminar participants at the ECB, the Universities of Glasgow, Brescia, Bari and Strathclyde. The usual disclaimer applies.

© European Central Bank, 2000

<b>Address</b>	<b>Kaiserstrasse 29 D-60311 Frankfurt am Main Germany</b>
<b>Postal address</b>	<b>Postfach 16 03 19 D-60066 Frankfurt am Main Germany</b>
<b>Telephone</b>	<b>+49 69 1344 0</b>
<b>Internet</b>	<b><a href="http://www.ecb.int">http://www.ecb.int</a></b>
<b>Fax</b>	<b>+49 69 1344 6000</b>
<b>Telex</b>	<b>411 144 ecb d</b>

*All rights reserved.*

*Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.*

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank.*

ISSN 1561-0810

# Contents

Abstract	5
<hr/>	
1 Introduction	7
<hr/>	
2 Investigating leading indicator properties of M3 by means of Granger causality tests within the Coenen and Vega (1999) framework.	9
<hr/>	
3 Leading indicator properties of M3 within an extended framework	11
3.1 Gerlach and Svensson (1999) P-star model of inflation in the euro area	11
3.2 Extending the analysis	12
3.3 What leading information in the real money gap?	16
3.4 On the robustness of the results	16
<hr/>	
4 Encompassing rival models	17
<hr/>	
5 Conclusions	20
<hr/>	
References	22
<hr/>	
Tables and figures	24
<hr/>	
European Central Bank Working Paper Series	32



## Abstract

The information content of broad money M3 for future GDP inflation in the euro area is investigated from a number of perspectives. Firstly, tests that money does not Granger-cause prices are conducted within a cointegrated VAR system comprising real M3 holdings, real GDP, inflation and short- and long-term interest rates. Secondly, this empirical framework is extended to investigate the claim that –in the context of an extended P-star model– the real money gap has substantial predictive power for future inflation. And thirdly, the P-star type of model developed is compared with an existing rival model of inflation in the euro area where no explicit role is given to monetary developments.

Our empirical results confirm that a significant positive association exists between the real money gap and future inflation up to five to six quarters ahead, reaching a maximum at the three-to-four quarter horizon. It is also shown that, although the extended P-star model outperforms the competing model in terms of out-of-sample forecast accuracy (as measured by the root mean square forecast errors) at horizons above two quarters, the hypothesis that no useful information is contained in rival evidence can be rejected at standard confidence levels.

Keywords: euro area, inflation, monetary aggregates, M3, leading indicators, P-star  
JEL Classification System: C32, C50, E30, E40





## I Introduction

In October and December 1998, the Governing Council of the European Central Bank announced the key elements of the monetary policy strategy of the Eurosystem. These comprise a quantitative definition of its primary policy objective, namely price stability, and *two pillars* which constitute the basis for its monetary policy strategy: a prominent role for monetary aggregates, as signalled by the announcement of a reference value for broad money M3, and a broadly based assessment of the outlook for, and risks to, price stability in the euro area<sup>1</sup>.

As regards the first pillar, there are good theoretical reasons -namely the claim that inflation is ultimately a monetary phenomenon - for underlying the *specialness* of monetary aggregates among the set of variables which are closely monitored by central banks. This relationship between money growth and inflation is present in almost every well-articulated economic model, thereby warranting, on its own, that money be given a prominent role.

At a more practical level, economic theory also suggests a number of possibilities for the implementation of this prominent role for money within a monetary policy strategy aimed at price stability. Firstly, monetary aggregates may play a nominal-anchor role, whereby the announcement of a reference for money growth well into the future help economic agents to form expectations about prices. This function entails quite a few practical requirements: i) a relationship linking money and prices must exist, be reasonably stable (predictable) and be invariant to central bank's own actions (i.e., be structural in the sense of being immune to the Lucas critique); ii) the announcement must be credible, implying commitment on the part of the central bank and requiring in turn that the money stock is controllable at some time horizon by means of the instruments available to the central bank (typically, a short-term interest rate). For the euro area, there exists some empirical evidence that broad money M3 may not be controllable at short-term horizons, precluding a pure monetary targeting strategy. However, available empirical evidence also points to both i) and ii) being fulfilled at medium-term horizons by M3<sup>2</sup>. This observation was indeed the main empirical basis for the selection of this aggregate when the first medium-term reference value for M3 money growth was announced in December 1998.

Money may also play a role within the strategy as an information variable to be included within the set of indicators which are regularly reviewed in the assessment of the outlook for, and risks to, price stability. This function entails one practical requirement, namely that past and current monetary developments contain valuable information about future prices, i.e. money being a leading indicator for inflation<sup>3</sup>.

From an empirical viewpoint, the investigation of leading indicator properties of money can be conducted following two different (but not incompatible) strategies. On the one hand, a structural model of the economy can be estimated and, subsequently, the indicator properties of money can be directly inferred from the restricted reduced form of the model. Alternatively, the indicator properties of monetary aggregates can be investigated straight from the unrestricted reduced form of the model itself by means, for instance, of well-known Granger non-causality tests. The latter is the most common approach in the literature but a number of serious concerns need to be born in mind.

<sup>1</sup> See ECB (1999). Angeloni et al. (1999) discuss the analytical foundations of the monetary policy strategy of the ECB.

<sup>2</sup> See Coenen and Vega (1999) and the references therein.

<sup>3</sup> At the conceptual level, the two roles for money –nominal anchor and leading indicator– need to be distinguished. The first is about whether knowledge of **future** monetary developments (the central bank announcement) help economic agents to predict inflation. The second is about whether **past** monetary developments help predict inflation, which is neither necessary nor sufficient for the first. Indeed, as shown in Sims (1972), the necessary and sufficient condition for the first is the existence of Granger causality from prices to money. This direction of causality is implied by the results reported in Coenen and Vega (1999). Also note that this view contrasts with the starting point in Estrella and Mishkin (1997), who maintain that in all potential uses of monetary aggregates “it is necessary first and foremost that the aggregates have some values as information variables”.

First, the concept of Granger causality refers to the entire universe of information (the complete vector of variables which characterises an economy) and, therefore, is intrinsically not operational. On the other hand, practical applications are necessarily confined to limited information sets (sometimes quite limited due to the lack of degrees of freedom frequently available to the applied researcher). In this context, there is no guarantee that inference does not change with the information set under consideration<sup>4</sup>. This observation strongly warns against conclusions drawn from reduced forms that are unlikely to be derived from a minimally well-specified structural model. In this respect, Woodford (1994) noted -as Granger (1988) did too - that this problem may be particularly acute when dealing with policy related variables: "...the mere fact that an indicator is found not to enter significantly in an inflation forecasting regression does not necessarily mean that " the central bank "...should be advised not to pay any attention to that variable. For the absence of forecasting power might simply imply that the variable is already being used..." by the central bank "in making policy, and in approximately the right way..."

And second, identifying which specific characteristics of the transmission mechanism account for stylised facts documented via the estimation of a reduced form is most crucial if any empirical finding concerning indicator properties of money were to be used in policy. Otherwise, absent structural interpretations, one would need to conclude with Woodford (1994): "...the mere fact that an indicator is found to be useful in forecasting inflation does not tell much about the desirability of a policy that involves feedback from that indicator...the ability of the indicator to signal the underlying sources of inflationary pressures that one wants to respond to may be impaired by the very fact that the monetary authority responds to it".

With these limitations in mind, the paper proceeds as follows. First, in **Section 2** leading indicator properties of broad money M3 are analysed by conducting standard tests that money does not Granger-cause prices in the context of a cointegrated VAR system comprising real holdings of M3 ( $m-p$ ), the inflation rate (as measured by the annualised quarterly changes in the GDP deflator,  $\pi=4\Delta p$ ), real GDP ( $y$ ) and the short- ( $R^s$ ) and long-term ( $R^l$ ) interest rates. This system was estimated in Coenen and Vega (1999) to investigate M3 money demand in the euro area. In **Section 3** we extend this empirical framework to look into the recent claim in Gerlach and Svensson (1999) that - in the context of an extended P-star model of inflation - the real money gap (i.e., the gap between current real balances and long-run equilibrium real balances) has substantial predictive power for future inflation in the euro area. Finally, in **Section 4** we conduct a number of encompassing tests to compare the P-star type of model developed in the previous section with a competing structural model of inflation like the one contained in Fagan, Henry and Mestre (2000), where no explicit role is given to monetary developments. In **Section 5**, some tentative conclusions are drawn.

Our results illustrate the claim above that inference about Granger non-causality may vary with the information set at hand and confirm the findings in Gerlach and Svensson (1999) that - in the context of an extended P-star model of inflation - a significant positive association exists between the real money gap and future inflation up to five to six quarters ahead, reaching a maximum at the three-to-four quarter horizon. It is also shown that, although the extended P-star model outperforms the Fagan, Henry and Mestre (2000) model in terms of out-of-sample forecast accuracy (as measured by the mean square forecast errors) at horizons above two quarters, the hypothesis that no useful information is contained in rival evidence can be rejected at standard confidence levels. We conclude that each model appears to have a strength of its own: both of them incorporate some information which turns out to be relevant to explain GDP inflation at different horizons; both of them, however, also fail to provide, on an individual basis, a complete account of inflation developments in the euro area.

4 By contrast, cointegration is a property which is invariant to expansions of the information set.

## 2 Investigating leading indicator properties of M3 by means of Granger causality tests within the Coenen and Vega (1999) framework.

Following Coenen and Vega (1999) [CV henceforth], let  $Z_t = (m-p, \pi, y, R^s, R^l) = (m-p, \pi, X')$  be a vector of  $CI(1,1)$  variables consisting of real holdings of M3 ( $m-p$ ), the inflation rate (as measured by the annualised quarterly changes in the GDP deflator,  $\pi = 4\Delta p$ ), real GDP ( $y$ ) and the short- ( $R^s$ ) and long-term ( $R^l$ ) interest rates. Consider the following VAR model.

$$Z_t = \sum_{i=1}^k J^i Z_{t-i} + V_t \quad (2.1)$$

and its VECM representation:

$$\Delta Z_t = \sum_{i=1}^{k-1} A^i \Delta Z_{t-i} + \Gamma A' Z_{t-1} + V_t \quad (2.2)$$

with  $\Gamma$  and  $A$  ( $5, r$ ) full-rank matrices of loading coefficients and cointegrating relationships, respectively.

Model (2.2) with lag-length  $k=2$  and cointegration rank  $r=3$  was estimated in CV. In the present context, leading indicator properties of broad money M3 can be investigated by conducting tests that money does not Granger-cause prices, following the results in Toda and Phillips (1993, 1994). We define the following partitioned matrices:

$$Z_t = \begin{pmatrix} m_t - p_t \\ \pi_t \\ X_t \end{pmatrix}; \quad A^i = \begin{pmatrix} A_{11}^i & A_{12}^i & A_{13}^i \\ A_{21}^i & A_{22}^i & A_{23}^i \\ A_{31}^i & A_{32}^i & A_{33}^i \end{pmatrix} \quad i=1, \dots, k-1; \quad \Gamma = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \quad (2.3)$$

Toda and Phillips (1993, 1994) suggest some operational procedures for conducting Granger causality tests that are applicable in the case of testing the causal effects of one variable on another group of variables and viceversa. In particular, three sequential causality Wald-type tests are recommended:

$$(P1) \text{ Test } H_2^* \begin{cases} \text{if } H_2^* \text{ is rejected, test } H^* \\ \text{otherwise, test } H_\perp^* \end{cases}$$

$$(P2) \text{ Test } H_1^* \begin{cases} \text{if } H_1^* \text{ is rejected, test } H^* \\ \text{otherwise, test } H_\perp^* \end{cases}$$

$$(P3) \text{ Test } H_\perp^* \begin{cases} \text{if } H_\perp^* \text{ is rejected, reject the null} \\ \text{otherwise, test } H_2^* \text{ and } H_1^* \begin{cases} \text{if both are rejected, test} \\ H_{21}^* \text{ if } r > 1, \text{ or reject} \\ \text{the null if } r = 1 \\ \text{otherwise, accept the null} \\ \text{of noncausality} \end{cases} \end{cases}$$

where the different hypotheses are defined as follows (see Toda and Phillips (1993) for the asymptotic distributions of the corresponding Wald tests):

$$\begin{aligned}
H^* &: A_{21}^1 = \dots = A_{21}^{k-1} = 0 \text{ and } \Gamma_2 A_1' = 0; && \text{(short - and long - run causality)} \\
H_{\perp}^* &: A_{21}^1 = \dots = A_{21}^{k-1} = 0; && \text{(short - run causality)} \\
H_1^* &: A_1 = 0; && \text{(long - run exclusion)} \\
H_2^* &: \Gamma_2 = 0; && \text{(weak - exogeneity)} \\
H_{21}^* &: \Gamma_2 A_1' = 0; && \text{(long - run causality)}
\end{aligned}$$

In the context of the estimation of model (2.2) over the sample period 1980:Q1-1998:Q4, the following results are obtained (p-values in brackets):

$$\begin{aligned}
H^* &: \chi_2^2 = .251 (.882) \\
H_{\perp}^* &: \chi_1^2 = .069 (.793) \\
H_1^* &: \chi_3^2 = 12.72 (.005) \\
H_2^* &: \chi_3^2 = 26.43 (.000) \\
H_{21}^* &: \chi_1^2 = .251 (.617)
\end{aligned}$$

Therefore, on the basis of the Toda and Phillips tests conducted above, the null hypothesis that M3 money does not Granger-cause GDP prices within the information set at hand cannot be rejected at standard confidence levels. This inference appears to be robust to the choice of sample period. Hence, when the VAR is estimated recursively over the sample 1993:Q1 to 1998:Q4, the maximal test statistics for the hypotheses  $H^*$ ,  $H_{\perp}^*$ , and  $H_{21}^*$  are, respectively: 0.477, 0.268 and 0.937. All of them are well below the corresponding  $\chi^2$  critical values<sup>5</sup>. The same results hold if the estimation sample starts in 1985, dropping the first five years from the analysis.

One weakness of the Phillips-Toda tests conducted above is the fact that they rely on pre-tests for unit roots and cointegration. Accurate determination of the number of unit roots and cointegration rank in small samples may be problematic, leading to pre-test biases in Granger-causality tests conditioned on the estimation of these parameters<sup>6</sup>. Against this background, we investigate in the following whether the failure to reject Granger non-causality of  $m$  on  $p$  could be attributed to mis-specifications of the cointegration rank and/or the degree of integration of the variables comprising the CV money demand system.

To this end, we follow Toda and Yamamoto (1995)<sup>7</sup>. These authors show how VARs in levels can be estimated and general restrictions be tested even if the time series involved are integrated or cointegrated of an arbitrary order. Firstly, a usual lag selection procedure can be applied to determine the lag length ( $k^*$ ) in the VAR. Secondly, a  $(k^* + d_{\max})$ -th-order VAR can be estimated, where  $d_{\max}$  is the maximal order of integration suspected to occur among the variables in the system (typically 1 or 2). Thirdly, disregarding the last  $d_{\max}$  lagged terms, general restrictions on the first  $k^*$  coefficient matrices can be tested for using standard (asymptotic) inference.

In Table I (left panel) the suggestions in Toda and Yamamoto (1995) are used to test for Granger non-causality of  $m$  on  $p$  in VARs comprising two set of variables  $Z^1 = (m, p, y, R^s, R^l)$  and  $Z^2 = (m-p, \Delta p,$

5 It should be noted that the use of  $\chi^2$  critical values in this recursive context biases results towards rejection of non-causality, since no allowance is made for the endogenous search. In this sense, the conservative nature of the exercise should provide a strong indication that the conclusion on non-Granger causality is firmly grounded.

6 See Yamada and Toda (1998).

7 See also Dolado and Lütkepohl (1996).

$y, R^s, R^l$ ). The maximal order of integration is presumed to be 2 (I(2) model) or 1 (I(1) model) in  $Z^1$  and 1 (I(1) model) in  $Z^2$ , allowing for the possibility of nominal variables being I(2) but constraining real variables to be, at most, I(1). This is consistent with the findings in CV but encompasses models in which both nominal and real variables are assumed to be I(1). Hence, columns 1 and 2 in table 1 refer to specification in  $Z^1=(m, p, y, R^s, R^l)$  under alternative hypotheses on the maximal degree of integration among the variables comprising the system:  $d_{\max}=2$  (col. 1) or  $d_{\max}=1$  (col. 2). Column 3 refers to specification in  $Z^2=(m-p, \Delta p, y, R^s, R^l)$  under the hypothesis  $d_{\max}=1$ . The right panel of table 1 reports tests for model reduction from M1 to M2; this is a test for long-run homogeneity of money and prices. Finally, consistent with the evidence from both sequential F-tests for lag exclusion and information criteria, the VAR's lag length are set to  $k=2$  in the shaded cells in the table. However, possible mis-specifications of  $k$  are accounted for by reporting also results for  $k^*+1$  and  $k^*-1$ . The complete set of results in the table should permit to assess the robustness of the conclusions based on the Phillips-Toda tests conducted above to different assumptions on maximal order of integration ( $d_{\max}$ ) and the VAR lag-length ( $k^*$ ).

Table 2 reproduces the analysis for different subsets of the variables in the systems: a)  $Z_a^1=(m, p)$  and  $Z_a^2=(m-p, \Delta p)$ ; b)  $Z_b^1=(m, p, y)$  and  $Z_b^2=(m-p, \Delta p, y)$ ; c)  $Z_c^1=(m, p, y, R^s)$  and  $Z_c^2=(m-p, \Delta p, y, R^s)$ ; and d)  $Z_d^1=(m, p, y, R^l, R^s)$  and  $Z_d^2=(m-p, \Delta p, y, R^l, R^s)$ . The structure of the table is the same as that of table 1 except for the inclusion of the corresponding Phillips-Toda tests in column 4.

A number of conclusions can be drawn from the analysis in this section. Firstly, there is little empirical evidence for rejecting Granger non-causality of  $m$  on  $p$  within the information set used in CV. Both the Phillips-Toda tests and the Toda-Yamamoto tests fail to reject Granger non-causality of  $m$  on  $p$  at standard confidence levels. Secondly, this conclusion, which is stable throughout the sample, appears to be robust to a number of plausible assumptions concerning: i) the maximal degree of integration of the variables in the system; ii) the lag-length selected for the VAR; and, iii) the imposition or not of the long-run homogeneity of  $m$  and  $p$ , though the long-run homogeneity hypothesis cannot be rejected at standard confidence levels. Finally, the results in table 2 also suggest that the conclusion above is invariant to the reduction of the information set employed in CV (1999).

### 3 Leading indicator properties of M3 within an extended framework

#### 3.1 Gerlach and Svensson (1999) P-star model of inflation in the euro area

An attempt to analyse leading indicator properties of M3 within a more 'structural' framework is made in a recent paper by Gerlach and Svensson (1999) [GS henceforth]. They find that the real money gap (i.e., the gap between current real balances and long-run equilibrium real balances) has substantial predictive power for future inflation in the euro area. They conclude that the real money gap should be an important indicator for the Eurosystem. GS estimate the following P-star type model for the inflation rate<sup>8</sup>:

8 See Groeneweld (1998) and the references therein for an overview of the P-star model and for empirical evidence on its relevance for various euro area countries.

$$\pi_t = \pi_{t-1}^e + \alpha_m (\tilde{m}_{t-1} - \tilde{m}_{t-1}^*) + \alpha_y (y_{t-1} - y_{t-1}^*) + \varepsilon_t \quad (3.1a)$$

$$\pi_{t-1}^e = (1 - \alpha_\pi) \hat{\pi}_t + \alpha_\pi \pi_{t-1} \quad (3.1b)$$

$$\hat{\pi}_t = \exp(\tau_0 + \tau_1 t) \quad (3.1c)$$

$$\tilde{m}_t \equiv (m - p)_t = k_0 + k_y y_t - k_i (R_t^l - R_t^s) \quad (3.1d)$$

$$\Delta \tilde{m}_t = \delta_0 - \delta_1 [\tilde{m} - k_0 - k_y y + k_i (R^l - R^s)]_{t-1} - \delta_2 (\pi_t - \hat{\pi}_t) + \delta_3 \Delta \tilde{m}_{t-1} + \nu_t \quad (3.1e)$$

$$p_t^* \equiv m_t - k_0 - k_y y_t^* + k_i (R^l - R^s)^* \quad (3.1f)$$

$$\tilde{m}_t - \tilde{m}_t^* \equiv (m - p)_t - (m - p^*)_t \equiv -(p_t - p_t^*) = (m - p)_t - k^* - k_y y_t^* \quad k^* = [k_0 - k_i (R^l - R^s)^*] \quad (3.1g)$$

where  $\pi^e$ ,  $\hat{\pi}$ ,  $\tilde{m}$  and  $y^*$  stand, respectively, for inflation expectations, central bank's inflation objective, real money balances and potential output.

Equation (3.1a) nests the Phillips curve and price gap models of inflation. (3.1b) specifies how inflation expectations are formed. (3.1c) assumes that the monetary authority's inflation objective can be described within the available sample by a deterministic exponential trend. This equation implements at a practical level the belief that the decline in the inflation rate observed during the sample in the euro area is largely a reflection of the fall in the average (implicit) inflation objectives of national central banks due to their growing commitment to lower inflation. (3.1d) and (3.1e) are the postulated long- and short-run money demand functions. (3.1f) defines the equilibrium price level ( $p^*$ ) as the price level that would prevail for a given level of the money stock when the remaining variables are at their equilibrium levels. This is computed by inversion of the long-run money demand (3.1d). Finally, (3.1g) defines the real money gap as the negative of the price gap.

The long-run equilibrium of the model is in turn defined by the following conditions:

$$y_t = y_t^*; \quad p_t = p_t^* \quad (\tilde{m}_t = \tilde{m}_t^*); \quad (R_t^l - R_t^s) = (R^l - R^s)^*; \quad \pi_t = \pi_{t+1} = \hat{\pi}_{t+1} \quad (3.2)$$

GS estimate the model above in two stages. First, the long-run income elasticity of money demand ( $k_y$ ) is estimated using cointegration techniques. That provides for an estimate of the real money gap according to (3.1g). And second, equations (3.1a+3.1b), (3.1c) and (3.1e) are estimated using: i) OLS ; ii) weighted non-linear least squares system estimates; and, iii) system estimates imposing the constraint that the inflation objective is 1.5% in 1998:Q4. Focusing on the equation for  $p$ , the following estimates (standard errors in parentheses) resulted from iii)<sup>9</sup>:

$$\pi_t = .415 \hat{\pi}_t + .585 \pi_{t-1} + .163 (\tilde{m} - \tilde{m}^*)_{t-1} + .021 (y - y^*)_{t-1} + \hat{u}_t \quad (3.3)$$

(-)
(.089)
(.074)
(.094)

which can trivially be expressed as:

$$\pi_t - \pi_{t-1} \equiv \Delta \pi_t = - .415 (\pi_{t-1} - \hat{\pi}_t) + .163 (\tilde{m} - \tilde{m}^*)_{t-1} + .021 (y - y^*)_{t-1} + \hat{u}_t \quad (3.4)$$

Equation (3.4) provides empirical support for the so-called P-star model. It identifies three sources of dis-equilibrium which cause inflation to move from its current level: i) inflation expectations are not consistent with the inflation objective; ii) the price level is not at its equilibrium level; or, iii) output is not at its potential level<sup>10</sup>.

<sup>9</sup> See Table 1 in Gerlach and Svensson (1989).

<sup>10</sup> Although the output gap is not significant at standard confidence levels in (3.3), Gerlach and Svensson (1999) provide some evidence that the output gap may contain valuable information for predicting future information over and above that already contained in the real money gap.

### 3.2 Extending the analysis

In this section, we build heavily on the work by GS but introduce a number of modifications to their model to account for some of the empirical findings in CV.

In the context of the analysis in Section 2, the presence of  $(\tilde{m} - \tilde{m}^*)$  in (3.3) and (3.4) can be interpreted as implying that there is Granger-causality of money M3 on prices when the information set employed in CV is augmented to include potential output, on the one hand, and a measure of the monetary policy objectives, on the other<sup>11</sup>. As shown in Granger (1988), explicit consideration of the latter (imperfect it may be in the present context) may be particularly important when dealing with policy related variables in the context of testing for Granger non-causality. It also implies that weak exogeneity of prices for the parameters of long-run money demand can be rejected within the augmented information set. To see the latter, notice that the money gap  $(\tilde{m} - \tilde{m}^*)$  and the error correction term given by deviations of real money holdings from the long-run money demand (3.4) [ $ecm_t \equiv \tilde{m}_t - k_0 - k_y y_t + k_i (R_t^l - R_t^s)$ ] are trivially linked by the following identity:

$$ecm_t \equiv (\tilde{m} - \tilde{m}^*)_t - k_y (y - y^*)_t + k_i [(R^l - R^s)_t - (R^l - R^s)^*]$$

The interpretation above would call for augmenting the system estimated in CV to include both a measure of the monetary authority's inflation objective and potential output. However, extension of the Johansen-type of analysis is hardly feasible in practice (even if these extra variables were assumed to be strongly exogenous for the parameters of interest). Instead, we proceed as follows throughout this section.

First, we consider the system estimates of the three cointegration relationships found in CV within the application of the Johansen procedure to the VAR model consisting of real money holdings, real income, short- and long term interest rates, inflation and an unrestricted constant. These system estimates do not depend to any extent on weak exogeneity assumptions, though they are not far from those obtained in CV from the long-run solution of a conditional ADL model. For the sample up to 1998:Q4, the results are:

$$\begin{aligned} ecm1_t &\equiv (m - p)_t - 1.158 y_t + 1.278 \pi_t - c \\ &\quad (.037) \quad (.189) \\ ecm2_t &\equiv (R^l - R^s)_t - (R^l - R^s)^* \\ ecm3_t &\equiv (R^s - \pi)_t - (R^s - \pi)^* \end{aligned}$$

As noted in CV, whilst the parameter measuring the interest rate semi-elasticity of long-run money demand cannot be identified without further restrictions from the system analysis, the first cointegrating vector above permits to identify unambiguously the long-run income elasticity and - assuming that both interest rates enter the long-run money demand as a spread- the semi-elasticity of money demand with respect to inflation. Estimates of these two parameters are needed for the calculation of the real money gap when the inflation rate is added to the r.h.s. variables of (3.1d), following the specification for long-run money demand in CV<sup>12</sup>.

<sup>11</sup> It may be worth noting that the real money gap can be trivially decomposed as  $(m - \tilde{m})_t \equiv (m - \tilde{m})_{t-1} + \Delta(m - \tilde{m})_t$ . Therefore, (3.3) or (3.4) may be re-parameterised in terms of the two r.h.s variables in the identity, the second of which could in turn be interpreted - according to the analysis in Gerlach and Svenson (1999) - as the deviations of money growth from the reference value. It follows that, provided due account is taken of the prevailing liquidity situation as measured by the level gap  $(m - \tilde{m})_{t-1}$ , the information content of  $\Delta(m - \tilde{m})_t$  is identical to that of  $(m - \tilde{m})_t$ , contradicting the conclusion in GS that deviations of money growth from the reference value has limited predictive content for future inflation. A valid claim, however, is that, when interpreting monetary developments in relation to the reference value, one needs necessarily to take into account the prevailing liquidity situation, since money growth above/below the reference value may well coexist with negative/positive real money gaps.

<sup>12</sup> See section 2.1 in Coenen and Vega (1999) for a discussion of the specification of long-run money demand.

Next, following a general-to-specific approach, we estimate a parsimonious I(0) VAR for  $\Delta Z_t$  conditional on  $\Delta Z_{t-1}$ ,  $ecm1_{t-1}$ ,  $ecm2_{t-1}$ ,  $ecm3_{t-1}$ . Focusing on the equation for  $\pi$ , the resulting estimates (after less significant variables are excluded) are as follows:

$$\Delta\pi_t = .45 [(R^l - R^s) - (R^l - R^s)^*]_{t-1} + .54 [(R^s - \pi) - (R^s - \pi)^*]_{t-1} + \hat{v}_t \quad (3.5)$$

$T=73$ (1980:Q4-1998:Q4)	$R^2=0.30$	$\sigma=.88\%$	$DW=1.95$
$LM(1)=1.34$ (.252)	$LM(4)=1.46$ (.231)	$LM(1,4)=0.60$ (.667)	
$ARCH(4)=0.50$ (.734)	$HET=0.80$ (.553)	$NORM=0.01$ (.997)	
$RESET=1.33$ (.252)	$FOR(24)=7.36$ (.999)	$CHOW(24)=.31$ (.999)	

where  $(R^l - R^s)^*$  and  $(R^s - \pi)^*$  stand for the (constant) equilibrium spread and the (constant) equilibrium short term interest rate and where  $LM(i)$  and  $LM(1,i)$  are the Lagrange multiplier F-tests for residual autocorrelation of order  $i^{th}$  and up to the  $i^{th}$  order, respectively,  $ARCH$  is the Engle F-test for autoregressive conditional heteroskedasticity,  $HET$  is the White F-test for heteroskedasticity,  $NORM$  is the Doornik and Hansen  $\chi^2$ -test for normality,  $RESET$  is the regression specification F-test due to Ramsey,  $FOR$  and  $CHOW$  are the out-of-sample forecast test and the Chow test for parameter stability over the period 1993:Q1-1998:Q4. Figure 1a records the time series of fitted and actual values of  $\Delta\pi$ , the scaled residuals and the residual correlogram. Figure 1b shows a graphical summary of results from the recursive estimation of equation (3.5) over the 90's.

From a statistical point of view, the estimated equation appears well specified, with tests showing no signs of residual autocorrelation, heteroskedasticity or non-normality. The exclusion of additional dynamics ( $\Delta Z_{t-1}$  and  $ecm1_{t-1}$ ) from (3.5) is not rejected either [ $F(8,63)=0.88$  (.540)], so that  $\hat{v}_t$  is an innovation relative to the information set at hand. In particular, the test for excluding  $ecm1_{t-1}$  yields  $F(1,70)=0.25$  (.617). Furthermore, no major problems are detected when the equation is used for producing one-step ahead forecasts over the last six years of the sample. In the same vein, when estimated recursively, the parameters in (3.5) turns out to be relatively constant over the recent period once due account is taken of the marked reduction in the area-wide weighted average of ex-post real interest rates after exchange rate risk premia are no longer incorporated in national interest rates. This has been modelled by a step reduction in the equilibrium short-term real interest rate from 4.6% (approximately the weighted average of national ex-post real interest rates in the period 1980:Q1 to 1997:Q7) to 3.0% as from 1998Q1, four quarters ahead of the introduction of the single currency<sup>13</sup>. The equilibrium spread is estimated constant at around 0.60% (approximately its sample average value).

The specification in (3.5) is similar to that in Tzavalis and Wickens (1996) for the US<sup>14</sup>. It shows that both the spread and the (ex-post) real short-term interest rate help predict future GDP inflation in the euro area, though the low  $R^2$  suggests that little inflation variability is indeed explained. Moreover, as the coefficients of the spread and the short-term rate are not significantly different in (3.5), the equation can be re-parameterised in terms of the long-term real rate alone, indicating in turn that the most of the information about future inflation is contained in the level of

13 In Gerlach and Schnabel (1999) a measure of credibility adjusted equilibrium real interest rate for the euro area is obtained by regressing the weighted ex-post real rate on the average rate of depreciation of the nominal exchange rate against the Deutsche mark. The intercept of that regression could be claimed to be interpretable as the equilibrium real rate that would prevail assuming no depreciation vis-à-vis the DM. The resulting estimate is 3.55%, with a standard error of 0.96%.

14 Tzavalis and Wickens (1996) review the theoretical basis for using the term spread to forecast inflation. They show for the US that the Mishkin (1990) type of forecasting equation relating a measure of inflation n-periods ahead to the spread between the n-period and the one-period interest rates performs poorly when compared to a specification in which both the spread and the real one-period interest rate are included. They conclude that the real interest rate seems to contain far more information about future inflation than the slope of the yield curve. See also Alonso et al. (1997) for Spain.



the nominal long-term interest rate. This empirical finding would be consistent with relatively stable *ex-ante* real interest rates and small and/or not too volatile term and inflation risk premia.

As a final step, we augmented the parsimonious I(0) VAR with the additional terms  $(\pi_{t-1} - \hat{\pi}_t)$  and  $(y - y^*)_{t-1}$  appearing in GS specification (3.4), since -as shown above- the money gap  $(\tilde{m} - \tilde{m}^*)_{t-1}$  can be trivially expressed as a linear combination of the already included long-run relationships. For comparison purposes, we then map the I(0) system  $\Delta Z_t / \Delta Z_{t-1}, ecm1_{t-1}, ecm2_{t-1}, ecm3_{t-1}, (\pi_{t-1} - \hat{\pi}_t)$  and  $(y - y^*)_{t-1}$  into an I(0) system comprising the real money gap instead of  $ecm1_{t-1}$ . Focusing again on the equation for  $\pi$ , the general specification encompasses specifications (3.4) and (3.5) and can be written (omitting the terms  $\Delta Z_{t-1}$  for the sake of simplicity) as follows:

$$\Delta\pi_t = -\theta_0(\pi_{t-1} - \hat{\pi}_t) + \theta_1(\tilde{m} - \tilde{m}^*)_{t-1} + \theta_2(y - y^*)_{t-1} + \theta_3[(i^l - i^s) - (i^l - i^s)^*]_{t-1} + \theta_4[(i^s_{t-1} - \hat{\pi}_t) - (i^s - \pi)^*] + \hat{v}_t$$

where the real money gap is defined as:

$$(\tilde{m} - \tilde{m}^*)_{t-1} \equiv (m - p)_{t-1} - k^* - 1.158y^*_{t-1} + 1.278\hat{\pi}_t$$

with  $k^* = k_0 + k_1(R^L - R^S)^*$ , estimated using the average sample value.

Using the exponential trend postulated in GS for  $\hat{\pi}^{15}$  and calculating potential output by means of a Hodrick-Prescott<sup>16</sup> filter with smoothing parameter  $\lambda = 1600$ , estimation of (3.6) over the sample 1980:Q4-1998:Q4 provides the following results (standard errors in parentheses):

$$\Delta\pi_t = - .783 (\pi_{t-1} - \hat{\pi}_t) + .196 (\tilde{m} - \tilde{m}^*)_{t-1} + .262 (y - y^*)_{t-1} + .249 [(i^s_{t-1} - \hat{\pi}_t) - (i^s - \pi)^*] + \hat{v}_t \quad (3.7)$$

(.106)                      (.072)                      (.130)                      (.073)

T=73 (1980:Q4-1998:Q4)	R <sup>2</sup> =0.45	σ=.79%	DW=1.92
LM(1)=.72 (.399)	LM(4)=.85 (.359)	LM(1,4)=0.68 (.607)	
ARCH(4)=0.14 (.968)	HET=.221 (.986)	NORM=1.23 (.541)	
RESET=0.15 (.904)	FOR(24)=8.5 (.999)	CHOW(24)=.31 (.998)	

From a statistical point of view, the estimated equation appears once again well specified, with test statistics showing no signs of residual autocorrelation, heteroskedasticity or non-normality. The exclusion of additional dynamics ( $\Delta Z_{t-1}$  and  $ecm2_{t-1}$ ) from (3.7) is not rejected either [F(6,63)=1.31 (.315)]. Furthermore, no major problems are detected when the equation is used for producing one-step ahead forecasts over the last six years of the sample. When estimated recursively, the parameters in (3.7) turn out to be relatively constant over the recent period. Figure 2a records the time series of fitted and actual values of  $\Delta\pi$ , the scaled residuals and the residual correlogram. Figure 2b shows a graphical summary of results from the recursive estimation of the equation over the 90's.

Estimation results tend to give support to GS's claim that the real money gap has substantive predictive power for future inflation in the euro area, as evidenced by the significance of the term

15 This is here obtained by separate OLS estimation of (3.1c). As noted in Gerlach and Svensson (1999), this procedure does not guarantee that proper account is taken of measurement errors due to the unobservable nature of  $\hat{\pi}$ . Whilst this could have been tackled by further extending the system to include (3.1c), GS results indicate that the system estimates provide the implausible result that the implicit inflation objective is 0.1% at the end of the sample when no restrictions are imposed on (3.1c). On the other hand, when the restriction that the inflation objective at the end of the sample equals 1.5% is imposed, little appears to be gained from estimating (3.1c) jointly with the rest of equations in the system.

16 The use of the HP filter to proxy the output gap is popular in applied work [see, for instance, Roberts (1997) in the context of the estimation of a Phillips curve model or Taylor (1998) in the context of monetary policy rules] although it is certainly not exempt from criticism [see, for instance, King and Revelo (1993) or Harvey and Jaeger (1993)]. As before, the estimates herein do not take into account the existence of potential measurement errors, nor do they take into account the two-sided nature of the filter, which may cause violations of weak and strong exogeneity assumptions.

$(\tilde{m} - \tilde{m}^*)_{t-1}$  in (3.7). Contrary to the predictions of the pure P-star model, (3.7) shows that the real money gap (or the price gap) is not a sufficient statistic for future inflation. This is consistent with the analysis in GS, who augment the basic P-star model with additional terms pertaining to a traditional Phillips curve model of inflation giving rise to specification (3.1+3.2). Yet, the presence of additional terms in (3.7) indicates that the extended P-star model given by (3.1+3.2) is badly misspecified, though estimates do not provide any indication on whether the source of this misspecification comes from 3.1 -i.e., the postulated extended P-star specification itself- or, most likely, from 3.2 -i.e., the postulated model for expectations formation-. Investigating what structural model may give rise to the stylised facts encountered is far above the modest objectives set for this paper and will need to be addressed in future research.

### 3.3 What leading information in the real money gap?

In this section we further investigate the main conclusion in Section 3.2 (namely that the real money gap has substantive predictive power for future inflation in the euro area) by analysing the correlations at different time horizons between  $(\pi - \hat{\pi})_t$  and the various indicators included in (3.7). Results in this respect are reported in Table 3. More precisely, the table shows estimates of equation (3.8) below for forecasting horizons ranging from  $h=1$  to  $h=8$  quarters:

$$(\pi - \hat{\pi})_{t+h-1} = \alpha_1^h (\pi_{t-1} - \hat{\pi}_t) + \alpha_2^h (\tilde{m} - \tilde{m}^*)_{t-1} + \alpha_3^h (y - y^*)_{t-1} + \alpha_4^h [(i_{t-1}^s - \hat{\pi}_t) - (i^s - \pi^*)] + \hat{v}_{t+h-1} \quad (3.8)$$

with  $v_{t+h-1}$  following a MA(h-1) process<sup>17</sup>.

The results in Table 3 highlight the significant positive association that exists between the real money gap and future inflation. This significant correlation appears to extend over five to six quarters and reaches a maximum at the three-to-four quarter horizon. Consistent with the findings in Section 3.2, the real money gap does not come out of the analysis in the Table as a sufficient statistic for future inflation. The other two gaps included in (3.7) - the output gap and the difference between the real short-term rate and the estimated equilibrium real short-term rate - also appear to contain valuable information over and above that already contained in the real money gap. According to the reported estimates, the correlation between the output gap and future inflation is indeed higher than in the case of the real money gap. It extends over six quarters and reaches a maximum at the five-quarter horizon. The leading information contained in the real rates is short-lived, extending over just three quarters. Finally, results in the table also show how forecasting performance deteriorates dramatically as the time horizon increases, as evidenced by the sharp reduction in the  $R^2$  of the regressions. None of the indicators analysed appear to contain information that can be used to forecast inflation at horizons over 6 quarters.

**Figure 3** illustrates further the main results above. It shows time series for inflation deviations from "target" together with the various gap indicators entering (3.8). The latter are lagged according to the evidence in Table 3 to allow for a better visual impression of the correlations found.

### 3.4 On the robustness of the results

The empirical results so far obtained illustrate that leading indicator properties of M3 can be detected by augmenting the information set in CV to include potential output and a measure of the monetary authorities' (implicit) inflation objectives. In this extended framework, deviations

<sup>17</sup> It should be noted that for  $h=1$ , (3.8) is just a trivial re-parameterisation of equation (3.7). Equation (3.8) is often interpreted as a forecasting equation. See, for instance, Stock and Watson (1999) and Clements and Hendry (1996).

from "trend" of GDP inflation are explained as a function of the deviations of real money, real output and real short-term interest rates from their unobserved equilibrium values. All these variables involve measurement problems that we have somewhat obviated by using some ad-hoc empirical measures. This certainly introduces important caveats to the conclusions reached in the previous sections and, moreover, makes it necessary to get a sense of how sensible results are to changes in our empirical measure of the un-observables .

With a view to analysing the robustness of the conclusions, estimates of (3.7) are reported in Table 4 for various empirical measures of  $\hat{\pi}$  and  $y^*$ . In the left panel of the table,  $\hat{\pi}$  is estimated by OLS using the deterministic exponential trend postulated in GS, whilst a number of measures of potential output are alternatively employed, namely: the one contained in Fagan, Henry and Mestre (2000)<sup>18</sup> and several measures obtained by smoothing real GDP through the application of a standard Hodrick-Prescott (HP) filter with parameter  $\lambda=1600, 5000$  and  $15000$ , respectively. In the right panel, the exercise is repeated for fixed  $y^*$  (constructed using the HP filter with  $\lambda=1600$ ) and for various HP-based measures of  $\hat{\pi}$ .

A number of conclusions can be drawn from the estimates in Table 4:

- The real money gap enters significantly most of the specifications in the table, thereby confirming the leading indicator properties found in GS. Whilst estimates vary across specifications, this qualitative result appears to be robust to the different empirical measures of potential output employed. It seems also quite, but not totally, robust to the use of different empirical measures of the monetary authority's inflation objective.
- The significance of the output gap in (3.6) appears to vary crucially with its own empirical measurement. The smoother the filter employed for deriving potential output the less significant the output gap becomes in equation (3.6).
- The reported estimates also show that the real interest rates and (to a lesser extent) the term spread also contain useful information for future inflation over and above that already contained in the real money gap and the output gap. This finding also appears quite robust across specifications.

## 4 Encompassing rival models

In Section 3.2, we estimated a model for GDP inflation which accounted for the most salient features of both GS and CV specifications. The estimated equation (3.7) was thoroughly evaluated against a number of statistical criteria. These included testing for homoskedastic innovation errors and constant parameters as well as testing for no information losses as we (following a general-to-specific approach) move along the reduction process to arrive at our preferred specification. That, in turn, ensures that the estimated model is data congruent, in the sense that it matches the available evidence. Congruency is an important operational criterion that may help discriminate between empirical models since 'true' models may be expected to be congruent. However, a model which is congruent is not necessarily 'true'<sup>19</sup> and a number of congruent models can be developed in empirical work. Moreover, congruent models may not be nested, which complicates comparisons between competing models.

Following Mizon and Richard (1986) and Hendry and Richard (1989), we define a congruent encompassing model as a model which is congruent and which is able to account for, or explain,

<sup>18</sup> Potential output is estimated therein from a Cobb-Douglas production function with smoothed Solow residuals.

<sup>19</sup> For instance, a stochastic process may be an innovation relative to a given information set but may be predictable when the information set is augmented, as already shown in Section 3.2.

the results obtained by rival models. Encompassing and forecast encompassing -the criteria used throughout this section for model evaluation- are more stringent concepts than 'better fit' or (when applied in the context of forecasting) 'lower root mean square forecast error', since it entails that no useful additional information is contained in rival evidence as well as variance / root mean square forecast errors dominance<sup>20</sup>.

In the present context, one example of a non-nested congruent model for GDP inflation can be found in Fagan, Henry and Mestre (2000). In their model, GDP prices are determined in the long run by trend unit labor costs, which depend in turn on potential GDP and the NAIRU. In the short-run, GDP inflation is explained as a function of changes in trend unit labour costs, changes in import prices and deviations of real trend unit labor costs from equilibrium. Estimation of the Fagan, Henry and Mestre (2000) model over the sample 1980:Q4-1997:Q4 provides the following results:

$$\Delta\pi_t = .021 - .776 \pi_{t-1} + .140 \Delta w_t^* + .068 \Delta w_{t-1}^* + .136 \Delta w_{t-2}^* + .024 \Delta pm_{t-1} - .274 (p - w^* - \theta)_{t-1}$$

(.005) (.109) (.043) (.047) (.045) (.014) (.076) (4.1)

T=69 (1980:Q4-1997:Q4)	R <sup>2</sup> =0.52	σ=.78%	DW=2.02
LM(1)=0.22 (.641)	LM(4)=0.23 (.634)	LM(1,4)=0.54 (.709)	
ARCH(4)=0.48 (.750)	HET=0.38 (.964)	NORM=4.12 (.127)	
RESET=1.55 (.218)	FOR(24)=16.019 (.887)	CHOW(24,37)=0.535 (.945)	

where  $w^*$  and  $pm$  stand, respectively, for trend unit labour costs and import prices.

With a view to discriminating between the two contending models, Mizon and Richard (1986) simplification encompassing tests (SET) were conducted. Tests for (3.7) and (4.1), respectively, against the minimal nesting model provided the following results (p-values in brackets): F(7,58)= 1.518 (.179) when the null is (3.7) and F(4,58)= 0.599 (.665) when the null is (4.1). Therefore, based on within-sample information, there appears to be little to choose between both models<sup>21</sup>: none of the test statistics can reject at standard confidence levels that both models (3.7) and (4.1) are valid simplifications from the minimal nesting model. This inference appears to be stable when the test statistics are computed recursively over the sample 1992:1-1997:4.

We also tested for out-of-sample forecast encompassing, with forecasts under models (3.7) and (4.1) generated using only lagged information. More precisely, the following equations were estimated recursively over the sample 1992:Q1-1997:Q4 to produce 24 observations of out-of-sample forecasts at horizons ranging from h=1 to h=8 quarters<sup>22</sup>:

$$(\pi_{t+h-1} - \pi_{t-1}) = \alpha_1^h (\pi_{t-1} - \hat{\pi}_t) + \alpha_2^h (\tilde{m} - \tilde{m}^*)_{t-1} + \alpha_3^h (y - y^*)_{t-1} + \alpha_4^h [(i_{t-1}^s - \hat{\pi}_t) - (i^s - \pi)^*] + \hat{\varepsilon}_{1+t-h-1} \quad (4.2)$$

$$(\pi_{t+h-1} - \pi_{t-1}) = \beta_0^h + \beta_1^h \pi_{t-1} + \sum_{i=1}^2 \beta_{2i}^h \Delta w_{t-i}^* + \sum_{i=1}^2 \beta_{3i}^h \Delta pm_{t-i} + \beta_4^h (p - w^* - \theta)_{t-1} + \hat{\varepsilon}_{2+t-h-1} \quad (4.3)$$

20 Ericsson (1992) reviews a number of measures of forecast performance, including encompassing, parameter constancy and mean square forecast errors. It shows that the two latter are necessary for 'good' forecast performance, but neither (nor both) is sufficient. See also Harvey et al. (1998).

21 Note also that the fits of the two models -as measured by their respective residual standard errors- are close.

22 In this exercise, the estimates of the different gap terms in (4.2) and (4.3) and, hence, the long-run parameters are based on end-of-sample information. Attempts to allow for recursive estimation of the long-run parameters fail to provide any systematic basis for comparison, giving rise to erratic results that proved to be difficult -if not impossible- to interpret. This happens despite the lack of convincing statistical evidence about parameter non-constancy, which is rejected by formal tests. Accordingly, we interpret this finding as a reflection of the problems involved in conducting inference on cointegration and in estimating long-run relationships using small samples, rather than a problem of genuine parameter instability. It should be kept in mind, however, that the proposed exercise downplays the uncertainty surrounding inflation forecasts which could have been produced using only the information truly available in real time, as demonstrated in Orphanides and van Norden (1999) in relation to the measurement of the output-gap.

Figure 4 shows the resulting inflation forecasts at horizons of one (top panel), four (middle panel) and eight (bottom panel) quarters. Visually, whilst (4.3) fares comparatively well at the one-quarter horizon, its performance deteriorates as the forecast horizon is extended. Conversely, (4.2) appears to outperform (4.3) at the four-quarter horizon, consistent with the finding above that the maximum correlation between the real money gap and future inflation is reached at that horizon. At the eight-quarter horizon, none of the models perform particularly well.

Table 5 below reports the corresponding root mean square forecast errors at the one- to eight-quarter forecast horizons together with some tests for forecasting encompassing<sup>23</sup>. When compared to the equation standard errors reported above, the results for  $h=1$  imply that both models fare particularly well during the period under analysis 1992:1-1997:4. In any event, Chow test statistics do not detect any structural break:  $\text{Chow}(24,41)=0.508$  (.960) in the case of (3.7) and  $\text{Chow}(24,37)=0.535$  (.945) in the case of (4.1). As regards forecasting encompassing, the table shows estimates of  $\alpha$  (with standard errors consistent to autocorrelation and heteroscedasticity) as well as p-values for the null hypotheses  $H_0: \alpha=0$  and  $H_1: \alpha=1$  in the regressions:

$$\varepsilon_{1t}^h = \alpha(\varepsilon_{1t}^h - \varepsilon_{2t}^h) + \zeta_t \quad (4.4)$$

where  $\varepsilon_1^h$  and  $\varepsilon_2^h$  ( $h=1, \dots, 8$ ) stand, respectively, for the corresponding forecast errors obtained from (4.2) and (4.3). Rejection of one of the nulls provides evidence in favour of the hypothesis that the rival model contains information which helps to explain forecast errors from the own model, i.e. (4.2) under  $H_0$  or (4.3) under  $H_1$ . This situation calls for forecast pooling, whereby better forecasts can be sought by combining the forecasts obtained from both models. As argued in Ericsson (1992), finding such a pooled forecast is *prima facie* evidence of both individual models being misspecified.

The evidence in the table points to model (4.2) outperforming model (4.3) at the one- and two-quarter horizons. That reveals that the P-star model may not be the best forecasting tool at short-term horizons<sup>24</sup>. Indeed, for  $h=1$  (4.2) appears to incorporate little information which is not already contained in (4.3). Conversely, the out-of-sample forecast record of the P-star model turns out to be comparatively better for all forecast horizons  $h \geq 3$ , though for  $h \geq 6$  performance is rather poor for both models. More importantly, however, the encompassing hypotheses are rejected at all horizons other than  $h=1$ . We interpret this outcome as an indication that, at those horizons, each model incorporates -though only partially- information which turns out to be relevant to explain inflation developments in the euro area.

Table 5					
h	$\alpha$	Model 4.2		Model 4.3	
		RMSE	$H_0: \alpha=0$	RMSE	$H_1: \alpha=1$
1	.830 (.348)	0.653	.0171**	0.585	.6251
2	.575 (.176)	0.656	.0011**	0.623	.0156**
3	.241 (.053)	0.638	.0000**	0.834	.0000**
4	.178 (.098)	0.614	.0701*	0.992	.0000**
5	.335 (.121)	0.827	.0058**	1.033	.0000**
6	.454 (.094)	1.094	.0000**	1.148	.0000**
7	.447 (.157)	1.145	.0043**	1.208	.0004**
8	.472 (.198)	1.259	.0173**	1.293	.0078**

23 See Chong and Hendry (1986) and Harvey et al. (1998).

24 Indeed, it could be argued that both models are of little practical use at very short horizons since there exist alternative non-causal statistical devices which typically fare pretty well at those horizons. For instance, if the interest is forecasting GDP inflation in a given quarter, one is likely to do better by using within-the-quarter monthly information in HICP or other price indices which are available ahead of the publication of GDP inflation.

## 5 Conclusions

The information content of M3 broad money for future inflation in the euro area was investigated from a number of perspectives throughout the paper. First, in **Section 2** the leading indicator properties of broad money M3 were analysed by conducting tests that money does not Granger-cause prices in the context of a cointegrated VAR system comprising real holdings of M3 ( $m-p$ ), the inflation rate (as measured by the annualised quarterly changes in the GDP deflator,  $\pi=4\Delta p$ ), real GDP ( $y$ ) and the short- ( $R^s$ ) and long-term ( $R^l$ ) interest rates. This system was estimated in Coenen and Vega (1999) for the purposes of investigating M3 money demand in the euro area. The main finding is that there appears to be little empirical support for rejecting at standard confidence levels Granger non-causality of  $m$  on  $p$  within this information set. This conclusion is found to be stable throughout the sample and seems to be robust to reductions of the information set and to a number of plausible assumptions concerning: i) the maximal degree of integration of the variables in the system; ii) the lag-length selected for the VAR; and, iii) the imposition or not of the long-run homogeneity of  $m$  and  $p$ , though the long-run homogeneity hypothesis cannot be rejected at standard confidence levels.

Noting that the information set in Coenen and Vega (1999) was not chosen to develop a model of inflation, we further investigated in **Section 3** leading indicator properties of broad money M3 by looking at the recent claim in Gerlach and Svensson (1999) that - in the context of an extended P-star model of inflation - the real money gap (i.e., the gap between current real balances and long-run equilibrium real balances) has substantial predictive power for future inflation in the euro area. We show that this empirical finding can be reproduced when the information set employed in Section 2 is augmented to include variables which enter Gerlach and Svensson (1999) P-star type model but which did not enter Coenen and Vega (1999) money demand system, namely potential output and the monetary authorities' implicit inflation objective. Whilst these variables are not directly observable and, therefore problems may certainly exist when it comes to measuring, it cannot be denied that they must enter any empirical model of inflation which is claimed to be theory consistent. The results in this section illustrate the claim made in the Introduction that inference about Granger non-causality may vary with the information set at hand and confirm that a significant positive association exists between the real money gap and future inflation up to five to six quarters ahead, reaching a maximum at the three-to four quarter horizon. This stylised fact appears to be quite, though not totally, robust to the use of a number of empirical measures for the two un-observables.

When it comes to providing structural interpretations of these findings, the results are however less satisfactory for the P-star model of inflation. We showed, also in Section 3, that the real interest rate and (to a lesser extent) the term spread appear to contain information that can be used to forecast inflation developments which are left unexplained by the Gerlach and Svensson (1999) extended P-star model for inflation. This in turn suggests that the treatment of inflation expectations in the GS model and most possibly the measurement of the monetary policy authorities implicit inflation objective, needs at least to be refined if one were to stick to the P-star type of proposed explanation. This, however, is left for future research.

Finally, in **Section 4** we compared the model developed in the previous section with an alternative model for GDP inflation in the euro area due to Fagan, Henry and Mestre (2000). In the latter, where no explicit role is given to monetary developments, GDP prices are determined in the long-run by trend unit labor costs, which depends in turn on potential GDP and the NAIRU. In the short-run, GDP inflation is explained as a function of changes in trend unit labour costs, changes in import prices and deviations of real trend unit labor cost from equilibrium. The comparison between the two models was made on the basis of their relative performance in explaining

inflation developments in the euro area, both within- and out-of-sample. The evidence in this section points to the P-star model outperforming the rival model at forecast horizons  $h \geq 3$  and being outperformed by the alternative model at shorter horizons. The encompassing hypothesis (i.e., the hypothesis that no useful information is contained in rival evidence) is, however, rejected at all horizons other than  $h=1$ . On the basis of this evidence, we conclude that each model appears to have a strength of its own: both of them incorporate some information which turns out to be relevant to explaining GDP inflation; both of them, however, also fail to provide on an individual basis a complete account of inflation developments in the euro area.

## References

**Alonso, F., Ayuso, J. and Pagés, J.** (1997): "El poder predictivo de los tipos de interés sobre la tasa de inflación española". Banco de España-Servicio de Estudios. Documento de Trabajo, 9722.

**Angeloni, I., Gaspar, V. and Tristani, O.** (1999). "The monetary policy strategy of the ECB", in D. Cobham and G. Zis (eds.), *From EMS to EMU*. London: Macmillan.

**Chong, Y. Y and Hendry, D.F.** (1986): "Econometric evaluation of linear macro-economic models". *Review of Economic Studies* 53, 671-690.

**Clements, M.P. and Hendry, D.F.** (1996): "Multi-step estimation for forecasting". *Oxford Bulletin of Economics and Statistics* 58 (4), 657-684.

**Coenen, G. and Vega, J.L** (1999): "The demand for M3 in the euro area". European Central Bank, WP #6.

**Dolado, J.J and Lütkepohl, H.** (1986): "Making Wald tests work in cointegrated VAR systems". *Econometric Review*, 15, 266-286.

**Ericsson, N. R** (1992): "Parameter constancy, mean square errors, and measuring forecast performance: an exposition, extensions and illustration". *Journal of Policy Modeling* 14 (4), 465-495.

**Estrella, A. and Mishkin, F.S.** (1997): "Is there a role for monetary aggregates in the conduct of monetary policy?". *Journal of Monetary Economics* 40, 279-304.

**European Central Bank** (1999). "The stability-oriented monetary policy strategy of the Eurosystem". *ECB Monthly Bulletin*, January, pp. 39-50.

**Fagan, G., Henry, J. and Mestre, R.** (2000): "An area-wide model (AWM) for the EU11". European Central Bank, forthcoming.

**Gerlach, S. and Schnabel** (1999): "The Taylor rule and interest rates in the EMU area: a note". Bank for International Settlements, WP #73.

**Gerlach, S. and Svensson, L.E.O.** (1999): "Money and inflation in the euro area: a case for monetary indicators?", mimeo.

**Granger, C.W.J.** (1988): "Causality, cointegration and control". *Journal of Economics Dynamics and Control* 12 (2-3).

**Groeneveld, J. M.** (1998): *Inflation patterns and monetary policy. Lessons for the European Central Bank*. Edward Elgar Pub. Ltd, UK and USA.

**Harvey, A.C and Jaeger, A.** (1993): "Detrending, stylised facts and the business cycle". *Journal of Applied Econometrics*, 8, 231-247.

**Harvey, D.I., Leybourne, S.J. and Newbold, P.** (1998): "Tests for forecast encompassing". *Journal of Business & Economic Statistics* 16 (2), 254-259.



**Hendry, D.F. and J.-F. Richard** (1989). "Recent developments in the theory of encompassing". In Cornet, B., and H. Tulkens (eds), *Contributions to operations research and econometrics*. The XX Anniversary of CORE, 393-440. Cambridge MA, MIT Press.

**King, R.G and Rebelo, S.** (1993): "Low frequency filtering and real business cycles". *Journal of Economic Dynamics and Control*, 17 (1-2), 207-231.

**Miskhin, F.S.** (1990): "The information in the longer-maturity term structure about future inflation". *Quarterly Journal of Economics* 25, 77-95.

**Mizon, G.E. and Richards, J.F** (1986): "The encompassing principle and its application to testing non-nested hypotheses". *Econometrica* 54, 657-678.

**Orphanides, A. and van Norden, S** (1999): "The reliability of output gap estimates in real time". Board of Governors of the Federal Reserve System. Finance and Economics Discussion Series

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

**Sims, C.A** (1972): "Money, income and causality". **American Economic Review**, 62, 540-552.

**Stock, J. H. and Watson, M. W** (1999): "Forecasting inflation". *Journal of Monetary Economics* 44, 293-335

**Taylor, J.B** (1998): "An historical analysis of monetary policy rules", NBER WP#6768.

**Toda, H.Y and Phillips, P.C.B** (1993): "Vector autoregression and causality". *Econometrica* 63, 1367-1393.

**Toda, H.Y and Phillips, P.C.B** (1994): "Vector autoregression and causality: a theoretical overview and simulation study". *Econometric Reviews* 13(2), 259-285.

**Toda, H.Y and Yamamoto, T.** (1995): "Statistical inference in vector autoregressions with possibly integrated processes". *Journal of Econometrics* 66, 225-250.

**Tzavalis, E and Wickens, M.R** (1996): "Forecasting inflation from the term structure". *Journal of Empirical Finance* 3, 103-122.

**Woodford. M.** (1994): "Nonstandard indicators for monetary policy: can their usefulness be judged from forecasting regressions?" In N.G Mankiw, ed. *Monetary Policy*, University of Chicago Press for NBER.

**Yamada, H. and Toda H.Y** (1998): "Inference in possibly integrated vector autoregressive models: some finite sample evidence". *Journal of Econometrics* 86, 55-95.

**Table 1**

	Tests for Granger noncausality of m on p in the VAR(k)			Wald tests for	
	M1:( m,p,y,s,l)		M2:( m-p,y,s,l, $\pi$ )	$H_0$ : M2 vs $H_a$ : M1	
	I(2) model $d_{\max}=2$ (1)	I(1) model $d_{\max}=1$ (2)	I(1) model $d_{\max}=1$ (3)	I(2) model $d_{\max}=2$ (4)	I(1) model $d_{\max}=1$ (5)
k=1	0.110 (0.740)	0.970 (0.325)	0.384 (0.535)	9.109 (0.105)	19.490 (0.109)
<b>k=2</b>	0.579 (0.749)	0.793 (0.673)	0.958 (0.619)	13.255 (0.151)	9.837 (0.364)
k=3	1.102 (0.777)	0.633 (0.889)	1.066 (0.785)	19.607 (0.106)	20.599 (0.245)

<b>Table 2</b>						
	<b>Toda and Yamamoto (1995) tests for Granger noncausality of m on p in the VAR(k)</b>			<b>Toda and Phillips tests for Granger noncausality of m on p</b>	<b>Wald tests for</b>	
	I(2) model $d_{\max}=2$	I(1) model $d_{\max}=1$	I(1) model $d_{\max}=1$		<b><math>H_0: M2</math> vs <math>H_a: M1</math></b>	
					I(2) model $d_{\max}=2$	I(1) model $d_{\max}=1$
	<b>M1:(m,p)</b>		<b>M2:(m-p,<math>\pi</math>)</b>	<b>M2:(m-p,<math>\pi</math>)</b> $r=0$		
k=1	0.004 (0.950)	0.134 (0.714)	0.425 (0.515)	$H^*=2.045$ (.153)	8.189** (0.017)	23.932*** (0.000)
<b>k=2</b>	<b>0.814</b> (0.666)	<b>1.410</b> (0.494)	<b>2.303</b> (0.316)		<b>5.105*</b> (0.078)	<b>6.737**</b> (0.034)
k=3	0.257 (0.968)	2.273 (0.518)	1.819 (0.611)		2.0599 (0.357)	0.059 (0.970)
	<b>M1:(m,p,y)</b>		<b>M2:(m-p,<math>\pi</math>,y)</b>	<b>M2:(m-p,<math>\pi</math>,y)</b> $r=1$		
k=1	0.097 (0.755)	0.006 (0.938)	0.614 (0.434)	$H_1^*=15.71$ (.000) $H_2^*=2.22$ (.136)	4.721 (0.193)	11.357** (0.010)
<b>k=2</b>	<b>0.518</b> (0.772)	<b>0.950</b> (0.622)	<b>1.033</b> (0.597)	$H_T^*=.345$ (.557)	<b>6.061</b> (0.109)	<b>15.633***</b> (0.001)
k=3	3.387 (0.336)	1.159 (0.763)	0.827 (0.843)	$H^*=3.969$ (.137)	2.150 (0.542)	1.690 (0.639)
	<b>M1:(m,p,y,s)</b>		<b>M2:(m-p,<math>\pi</math>,y,s)</b>	<b>M2:(m-p,<math>\pi</math>,y,s)</b> $r=2$		
k=1	0.970 (0.325)	2.470 (0.116)	1.337 (0.248)	$H_1^*=17.71$ (.000) $H_2^*=5.20$ (.074)	5.209 (0.267)	5.158 (0.272)
<b>k=2</b>	<b>0.846</b> (0.655)	<b>1.047</b> (0.592)	<b>0.420</b> (0.811)	$H_T^*=.013$ (.909)	<b>2.104</b> (0.717)	<b>2.699</b> (0.609)
k=3	2.104 (0.551)	1.046 (0.790)	0.381 (0.944)	$H^*=1.556$ (.459)	2.970 (0.563)	1.935 (0.748)
	<b>M1:(m,p,y,l-s)</b>		<b>M2:(m-p,<math>\pi</math>,y,l-s)</b>	<b>M2:(m-p,<math>\pi</math>,y,l-s)</b> $r=2$		
k=1	0.192 (0.661)	0.466 (0.495)	0.309 (0.578)	$H_1^*=22.72$ (.000) $H_2^*=2.59$ (.274)	4.601 (0.331)	4.215 (0.378)
<b>k=2</b>	<b>0.243</b> (0.886)	<b>0.335</b> (0.846)	<b>0.445</b> (0.801)	$H_T^*=.067$ (.796)	<b>4.726</b> (0.317)	<b>11.131**</b> (0.025)
k=3	0.312 (0.956)	0.810 (0.847)	0.527 (0.913)	$H^*=2.74$ (.253)	1.819 (0.769)	1.593 (0.810)

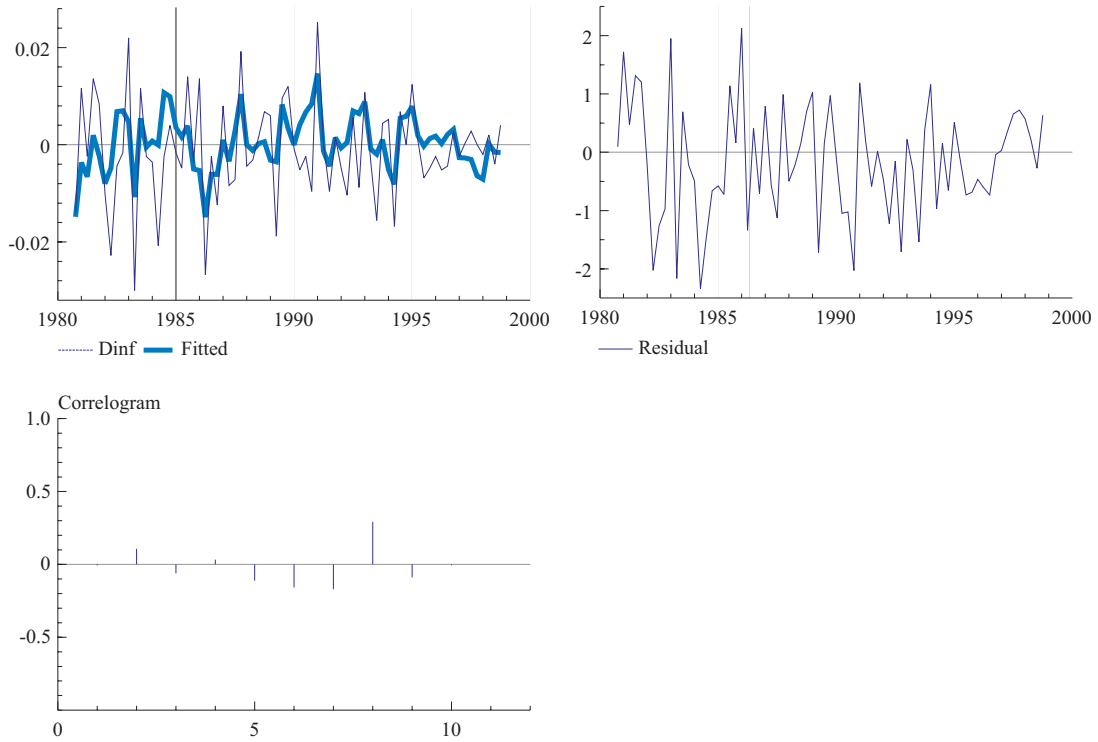
**Table 3**

	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9	h=10	h=11	h=12
$\hat{\alpha}_1^h$	.217**	-.012	-.124	-.053	-.147	-.159	-.114	.127	-.146	-.172	-.191	-.035
$\hat{\alpha}_2^h$	.196*	.283*	.355**	.341**	.316**	.205*	.028	-.119	-.076	-.076	-.087	-.206
$\hat{\alpha}_3^h$	.262**	.412**	.533**	.588**	.652**	.524**	.234	.096	.178	.240	.211	.161
$\hat{\alpha}_4^h$	.249*	.255*	.196**	.097	.113	.146	.211	.141	.212**	.149	.096	-.003
$\bar{R}^2$	.52	.45	.41	.33	.32	.21	.11	.11	.09	.06	.03	.07

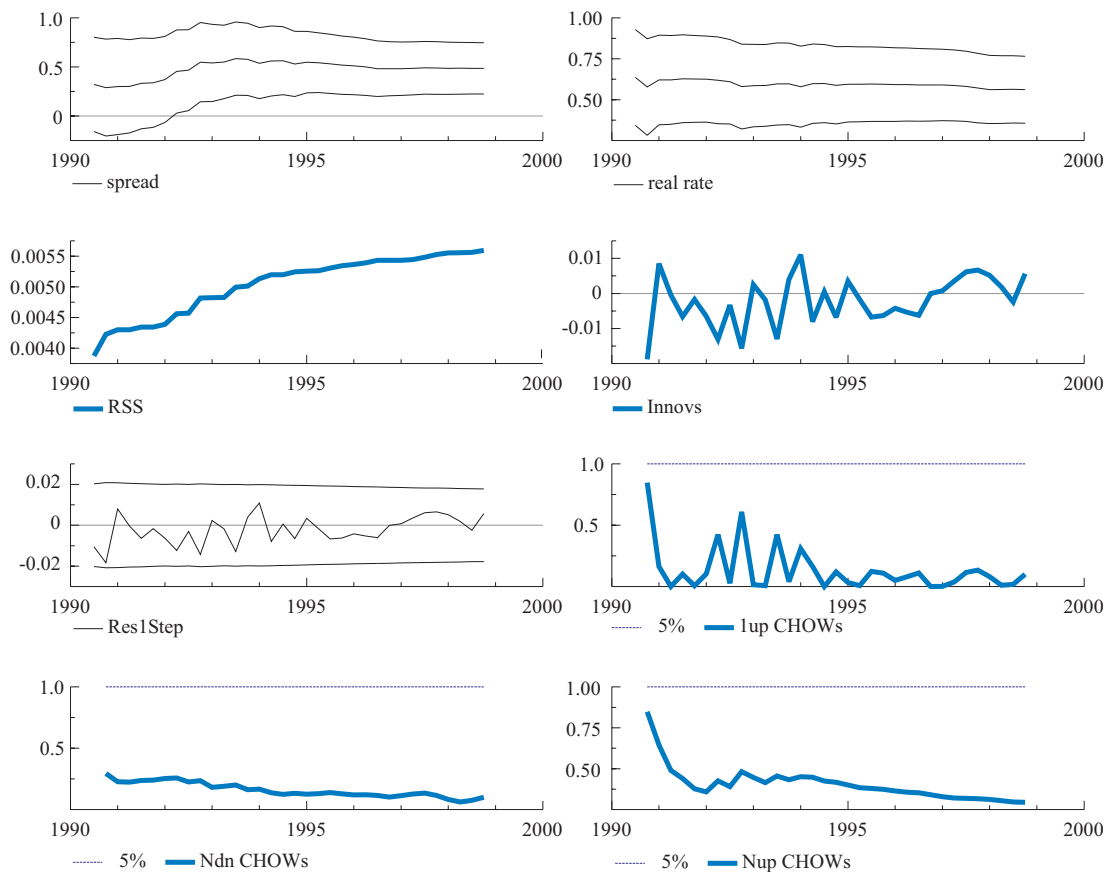
**Table 4**

	$\hat{\pi}$ = determ. exponent. trend				$y^* = \text{HP}(16000)$			
	$y^* = \text{AWM}$	$y^* = \text{HP}(1600)$	$y^* = \text{HP}(5000)$	$y^* = \text{HP}(15000)$	$\hat{\pi}$ = determ. exp. trend	$\hat{\pi}$ = HP(1600)	$\hat{\pi}$ = HP(5000)	$\hat{\pi}$ = HP(15000)
$\hat{\theta}_0$	<b>-0.774**</b> (.110)	<b>-0.812**</b> (.108)	<b>-0.811**</b> (.107)	<b>-0.809**</b> (.107)	<b>-0.812**</b> (.108)	<b>-0.831**</b> (.114)	<b>-0.827**</b> (.112)	<b>-0.814**</b> (.110)
$\hat{\theta}_1$	<b>.148**</b> (.068)	<b>.210**</b> (.076)	<b>.197**</b> (.070)	<b>.170**</b> (.065)	<b>.210**</b> (.076)	<b>.064</b> (.056)	<b>.109*</b> (.062)	<b>.175**</b> (.073)
$\hat{\theta}_2$	<b>.065</b> (.086)	<b>.264*</b> (.134)	<b>.154</b> (.096)	<b>.095</b> (.083)	<b>.264*</b> (.134)	<b>.195</b> (.132)	<b>.245*</b> (.136)	<b>.291**</b> (.141)
$\hat{\theta}_3$	<b>.340**</b> (.141)	<b>.211</b> (.133)	<b>.291**</b> (.132)	<b>.320**</b> (.134)	<b>.211</b> (.133)	<b>.226</b> (.136)	<b>.210</b> (.136)	<b>.188</b> (.138)
$\hat{\theta}_4$	<b>.446**</b> (.092)	<b>.348**</b> (.103)	<b>.362**</b> (.100)	<b>.366**</b> (.101)	<b>.348**</b> (.103)	<b>.271**</b> (.116)	<b>.282**</b> (.113)	<b>.281**</b> (.113)
	<b><math>\sigma = .80\%</math></b> <b><math>R^2 = .45</math></b>	<b><math>\sigma = .79\%</math></b> <b><math>R^2 = .47</math></b>	<b><math>\sigma = .78\%</math></b> <b><math>R^2 = .48</math></b>	<b><math>\sigma = .78\%</math></b> <b><math>R^2 = .48</math></b>	<b><math>\sigma = .79\%</math></b> <b><math>R^2 = .47</math></b>	<b><math>\sigma = .80\%</math></b> <b><math>R^2 = .45</math></b>	<b><math>\sigma = .80\%</math></b> <b><math>R^2 = .46</math></b>	<b><math>\sigma = .80\%</math></b> <b><math>R^2 = .46</math></b>

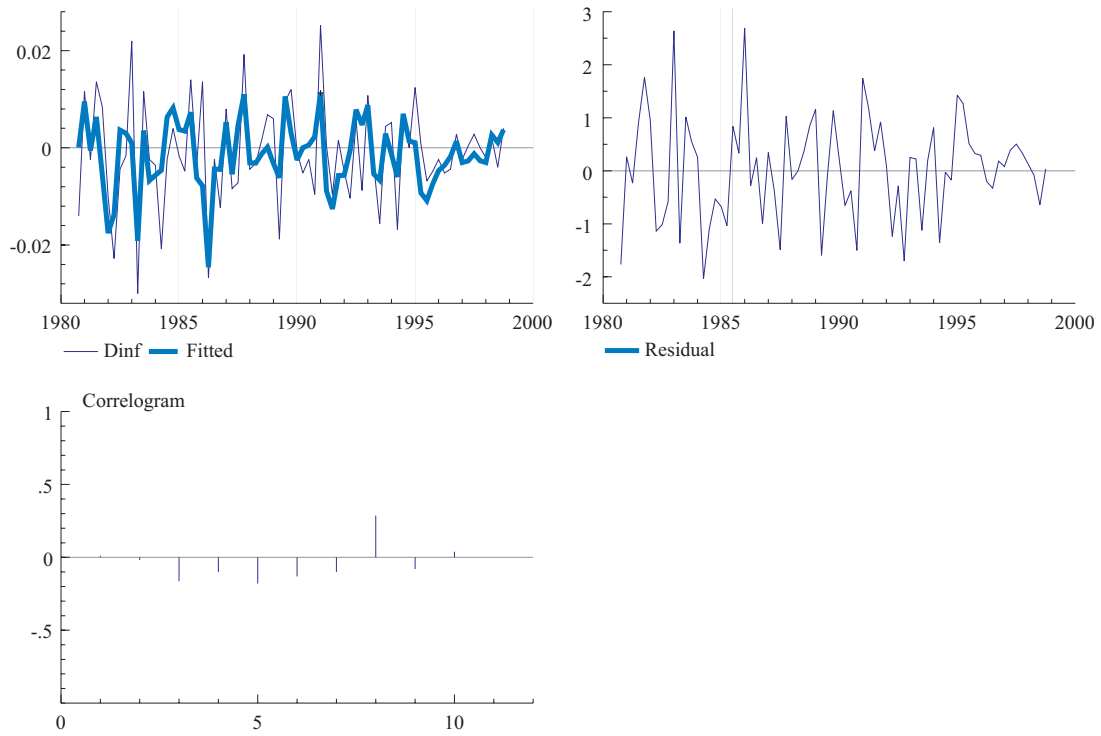
**Figure 1a**



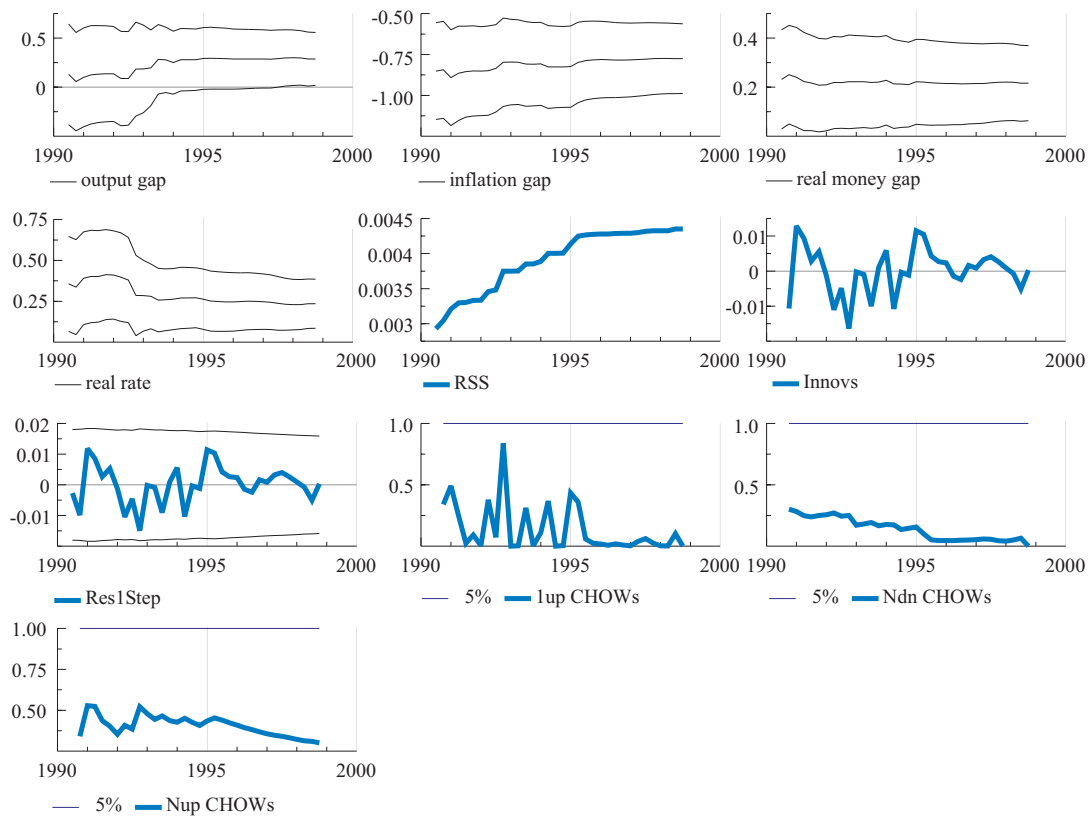
**Figure 1b**



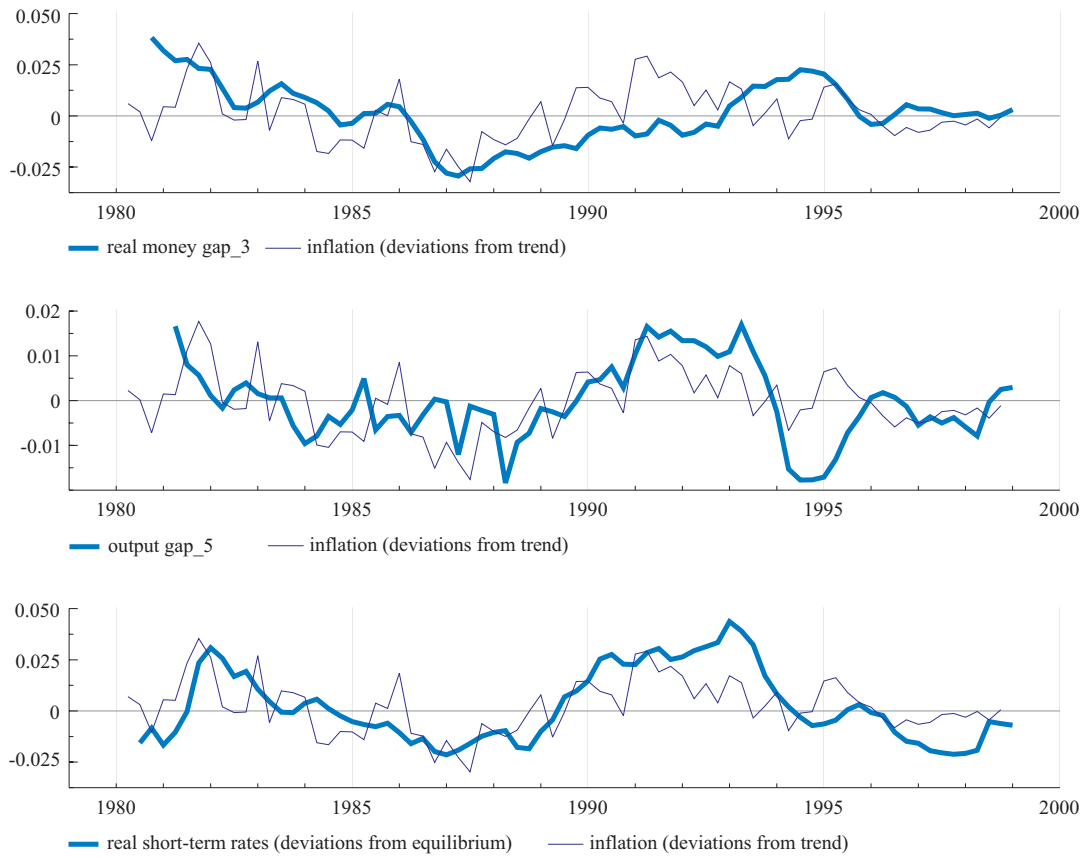
**Figure 2a**



**Figure 2b**



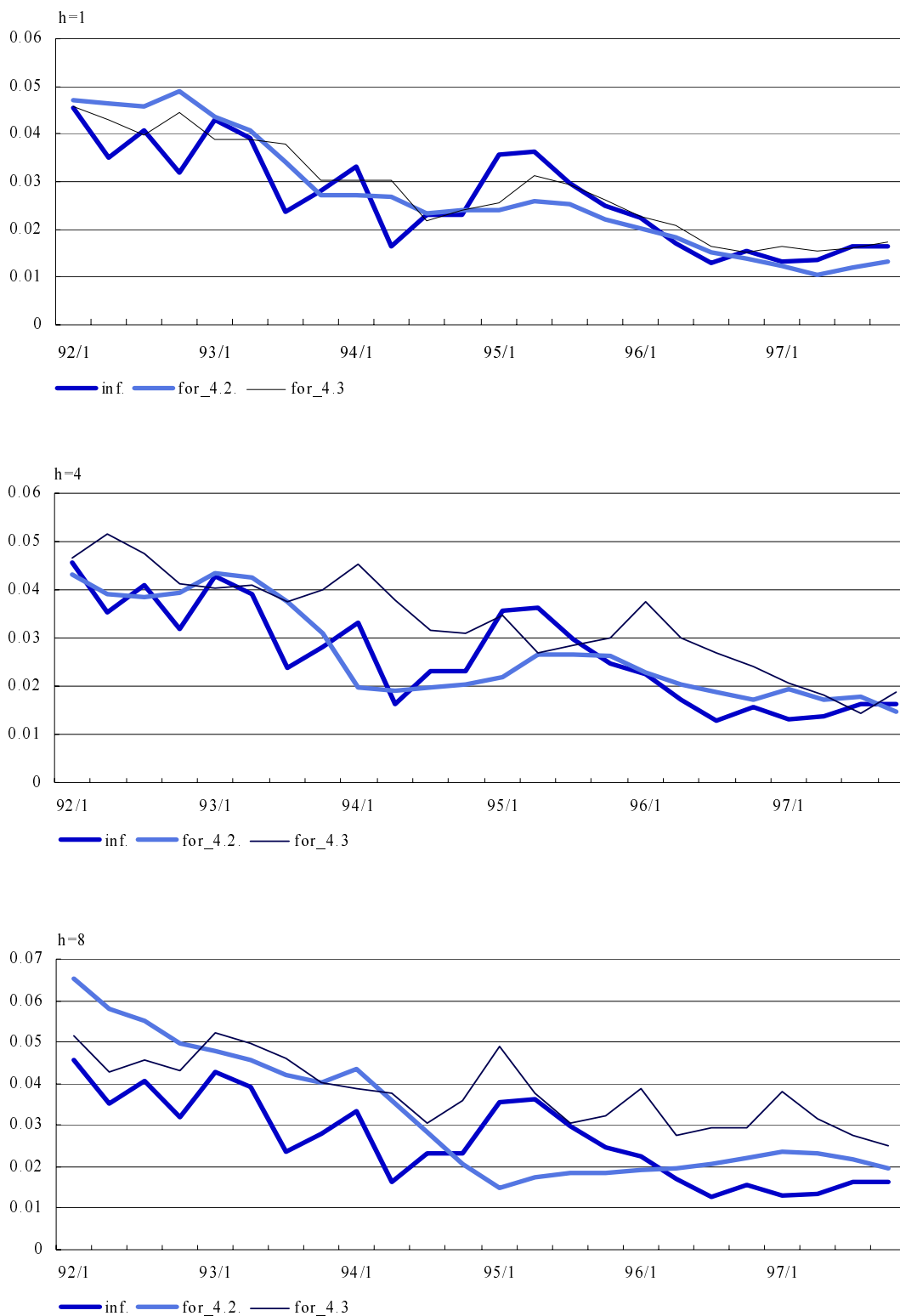
**Figure 3**





**Figure 4**

**Inflation forecasts at different horizons**



## European Central Bank Working Paper Series

- 1 “A global hazard index for the world foreign exchange markets” by V. Brousseau and F. Scacciavillani, May 1999.
- 2 “What does the single monetary policy do? A SVAR benchmark for the European Central Bank” by C. Monticelli and O. Tristani, May 1999.
- 3 “Fiscal policy effectiveness and neutrality results in a non-Ricardian world” by C. Detken, May 1999.
- 4 “From the ERM to the euro: new evidence on economic and policy convergence among EU countries” by I. Angeloni and L. Dedola, May 1999.
- 5 “Core inflation: a review of some conceptual issues” by M. Wynne, May 1999.
- 6 “The demand for M3 in the euro area” by G. Coenen and J.-L. Vega, September 1999.
- 7 “A cross-country comparison of market structures in European banking” by O. de Bandt and E. P. Davis, September 1999.
- 8 “Inflation zone targeting” by A. Orphanides and V. Wieland, October 1999.
- 9 “Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models” by G. Coenen, January 2000.
- 10 “On the effectiveness of sterilized foreign exchange intervention” by R. Fatum, February 2000.
- 11 “Is the yield curve a useful information variable for the Eurosystem?” by J. M. Berk and P. van Bergeijk, February 2000.
- 12 “Indicator variables for optimal policy” by L. E. O. Svensson and M. Woodford, February 2000.
- 13 “Monetary policy with uncertain parameters” by U. Söderström, February 2000.
- 14 “Assessing nominal income rules for monetary policy with model and data uncertainty” by G. D. Rudebusch, February 2000.
- 15 “The quest for prosperity without inflation” by Athanasios Orphanides, March 2000.
- 16 “Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model” by Peter Hördahl, March 2000.
- 17 “Alternative measures of the NAIRU in the euro area: estimates and assessment” by Silvia Fabiani and Ricardo Mestre, March 2000.
- 18 “House prices and the macroeconomy in Europe: Results from a structural VAR analysis” by Matteo Iacoviello, April 2000.

- 19 "The Euro and international capital markets" by Carsten Detken and Philipp Hartmann, April 2000.
- 20 "Convergence of fiscal policies in the euro area" by O. de Bandt and F. P. Mongelli, May 2000.
- 21 "Firm size and monetary policy transmission: evidence from German business survey data" by M. Ehrmann, May 2000.
- 22 "Regulating access to international large value payment systems" by C. Holthausen and T. Rønde, June 2000.
- 23 "Escaping Nash inflation" by I-K. Cho and T. J. Sargent, June 2000.
- 24 "What horizon for price stability" by F. Smets, July 2000.
- 25 "Caution and conservatism in the making of monetary policy" by P. Schellekens, July 2000.
- 26 "Which kind of transparency? On the need for clarity in monetary policy-making" by B. Winkler, August 2000.
- 27 "This is what the US leading indicators lead" by M. Camacho and G. Perez-Quiros, August 2000.
- 28 "Learning, uncertainty and central bank activism in an economy with strategic interactions" by M. Ellison and N. Valla, August 2000.
- 29 "The sources of unemployment fluctuations: an empirical application to the Italian case" by S. Fabiani, A. Locarno, G. Oneto and P. Sestito, September 2000.
- 30 "A small estimated euro area model with rational expectations and nominal rigidities" by G. Coenen and V. Wieland, September 2000.
- 31 "The disappearing tax base: is foreign direct investment eroding corporate income taxes?" by R. Gropp and K. Kostial, September 2000.
- 32 "Can indeterminacy explain the short-run non-neutrality of money?" by F. De Fiore, September 2000.
- 33 "The information content of M3 for future inflation" by C. Trecroci and J. L. Vega, October 2000.