The Information Content of M3 for Future Inflation in the Euro Area

By

Carmine Trecroci and Juan Luis Vega

Contents: I. Introduction. – II. Investigating the Leading Indicator Properties of M3 in a Money Demand VAR System. – III. Leading Indicator Properties of M3 within an Extended Framework. – IV. Encompassing Rival Models. – V. Conclusions.

I. Introduction

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 in the conduct of monetary policy.

At a more practical level, economic theory also suggests a number of possibilities for implementing a 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 between 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 (see Coenen and Vega 2001 and the references therein). 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 by the Governing Council of the European Central Bank in December 1998.[1]

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.[2]

From an empirical viewpoint, the investigation of leading indicator properties of money can be conducted by 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 borne in mind.

First, the concept of Granger causality refers to the entire universe of information (the complete vector of variables which characterizes an

---

[1] See ECB (1999). Issing et al. (2001) discuss the analytical foundations of the monetary policy strategy of the ECB.

[2] 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 (2001). 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".

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.[3] 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: 101) 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 characteristics of the transmission mechanism account for stylized 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, when structural interpretations are absent, one would need to conclude with Woodford (1994: 102): "...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 II leading indicator properties of broad money M3 are analyzed by conducting standard tests that money does not Granger-cause prices in the context of a cointegrated VAR system that has recently been proposed to investigate money demand in the euro area. In Section III, we extend this empirical framework to look into the recent claim that – in the context of a P-star model – 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 IV we conduct a number of encompassing tests to compare the model developed in the previous section with

---

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

a competing structural model of inflation where no explicit role is given to monetary developments. In Section V, 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 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 structural model in terms of out-of-sample forecast accuracy 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 strengths 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.

II. Investigating the Leading Indicator Properties of M3 in a Money Demand VAR System

Following Coenen and Vega (2001) [CV henceforth], let $Z' = (m - p, \pi, y, R^s, R^l) = (m - p, \pi, X')$ be a vector of CI(1,1) variables consisting of real holdings of M3, $(m - p)$, the inflation rate (as measured by the annualized 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:

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

and its VECM representation:

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

with Γ 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, 1994a, 1994b). We define the following partitioned matrices:

$$
\mathbf{Z}_t = \begin{pmatrix} m_t - p_t \\ \pi_t \\ X_t \end{pmatrix}; \quad \mathbf{A}^{i)} = \begin{pmatrix} A_{11}^{i)} & A_{12}^{i)} & A_{13}^{i)} \\ A_{21}^{i)} & A_{22}^{i)} & A_{23}^{i)} \\ A_{31}^{i)} & A_{32}^{i)} & A_{33}^{i)} \end{pmatrix} i = 1, \ldots, k-1;
$$

$$
\mathcal{G} = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}. \tag{2.3}
$$

Toda and Phillips (1993,1994a, 1994b) 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 vice versa. In particular, three sequential causality Wald-type tests are recommended:

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

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

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

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

H^*: $A_{21}^1 = \ldots = A_{21}^{k-1} = 0$ and $\Gamma_2 A_1' = 0$ (short- and long-run causality)
$H_\perp^*$: $A_{21}^1 = \ldots = A_{21}^{k-1} = 0$ (short-run causality)
H_1^*: $A_1 = 0$ (long-run exclusion)
H_2^*: $\Gamma_2 = 0$ (weak exogeneity)
H_{21}^*: $\Gamma_2 A_1' = 0$ (long-run causality).

The estimation of model (2.2) over the sample period 1980:Q1-1998:Q4 yields the following results (p-values in brackets): H^*: $\chi^2_2 = 0.251$ (0.882); $H^*_\perp$: $\chi^2_1 = 0.069$ (0.793); $H^*_\uparrow$: $\chi^2_3 = 12.72$ (0.005); H^*_2: $\chi^2_3 = 26.43$ (0.000); H^*_{21}: $\chi^2_1 = 0.251$ (0.617).

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 chosen sample period. When the VAR is estimated recursively over the sample 1993:Q1 – 1998:Q4, the maximal test statistics for the hypotheses H^*, $H_\perp$, and H^*_{21} are: 0.477, 0.268 and 0.937, respectively. All of them are well below the corresponding χ^2 critical values.[4] The same results hold when 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 (see Yamada and Toda 1998). 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 included in the CV money demand system.

To this end, we follow Toda and Yamamoto (1995)[5]. These authors show how levels VARs can be estimated and general restriction 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 1 (left panel) the suggestions in Toda and Yamamoto (1995) are used to test for Granger non-causality of m on p in VARs com-

---

[4] It should be noted that the use of χ^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.

[5] See also Dolado and Lütkhepohl (1996).

Table 1: *Tests for Granger Non-Causality of M3 on Prices in the VAR(k)*

| | Granger non-causality tests for | | | Wald tests for | |
| | M1: (m, p, y, s, l) | | M2: $(m-p, y, s, l, p)$ | H_0: M2 vs H_1: 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)
$k = 2$	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)

Note: The VARs include two set of variables: $Z^1 = (m, p, y, R^s, R^l)$ and $Z^2 = (m-p, \Delta p, y, R^s, R^l)$. 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. Hence, columns 1 and 2 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 reports tests for model reduction from M1 to M2. Finally, the VAR's lag length are set to $k = 2$ in the second row. However, possible mis-specifications of k are accounted for by reporting also results for $k^* + 1$ and $k^* - 1$.

prising two set of variables: $Z^1 = (m, p, y, R^s, R^l)$ and $Z^2 = (m - p, \Delta p, y, R^s, R^l)$. 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: $d_{max} = 2$ in col. 1 or $d_{max} = 1$ in 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 also set to $k = 2$. However, possible mis-specifications of k are accounted for by reporting also results for $k^* + 1$ and $k^* - 1$. The complete set of re-

Table 2: *Tests for Granger Non-Causality of M3 on Prices in the VAR(k) on a Subset of the Variables Employed for Table 1*

| | Tests for Granger non-causality | | | | Wald tests for H_0: M2 vs H_1: M1 | |
| | Toda and Yamamoto | | | Toda and Phillips | | |
	I(2) model $d_{max} = 2$	I(1) model $d_{max} = 1$	I(1) model $d_{max} = 1$		I(2) model $d_{max} = 2$	I(1) model $d_{max} = 1$
	M1: (m, p)		M2: $(m - p, \pi)$	M2: $(m - p, \pi)$ $r = 0$		
$k = 1$	0.004 (0.950)	0.134 (0.714)	0.425 (0.515)	$H^* = 2.045$ (0.153)	8.189** (0.017)	23.932*** (0.000)
$k = 2$	0.814 (0.666)	1.410 (0.494)	2.303 (0.316)		5.105* (0.078)	6.737** (0.034)
$k = 3$	0.257 (0.968)	2.273 (0.518)	1.819 (0.611)		2.0599 (0.357)	0.059 (0.970)
	M1: (m, p, y)		M2: $(m - p, \pi)$	M2: $(m - p, \pi, y)$ $r = 1$		
$k = 1$	0.097 (0.755)	0.006 (0.938)	0.614 (0.434)	$H_1^* = 15.71$ (0.000) $H_2^* = 2.22$ (0.136)	4.721 (0.193)	11.357** (0.010)
$k = 2$	0.518 (0.772)	0.950 (0.622)	1.033 (0.597)	$H_T^* = 0.345$ (0.557)	6.061 (0.109)	15.633*** (0.001)
$k = 3$	3.387 (0.336)	1.159 (0.763)	0.827 (0.843)	$H^* = 3.969$ (0.137)	2.150 (0.542)	1.690 (0.639)
	M1: (m, p, y, s)		M2: $(m - p, \pi, y, s)$	M2: $(m - p, \pi, y, s)$ $r = 2$		
$k = 1$	0.970 (0.325)	2.470 (0.116)	1.337 (0.248)	$H_1^* = 17.71$ (0.000) $H_2^* = 5.20$ (0.074)	5.209 (0.267)	5.158 (0.272)
$k = 2$	0.846 (0.655)	1.047 (0.592)	0.420 (0.811)	$H_T^* = 0.13$ (0.909)	2.104 (0.717)	2.669 (0.609)
$k = 3$	2.104 (0.551)	1.046 (0.790)	0.381 (0.944)	$H^* = 1.556$ (0.459)	2.970 (0.563)	1.935 (0.748)
	M1: $(m, p, y, l - s)$		M2: $(m - p, \pi, y, l - s)$	M2: $(m - p, \pi, y, l - s)$ $r = 2$		
$k = 1$	0.192 (0.661)	0.466 (0.495)	0.309 (0.578)	$H_1^* = 22.72$ (0.000) $H_2^* = 2.59$ (0.274)	4.601 (0.331)	4.215 (0.378)
$k = 2$	0.243 (0.886)	0.335 (0.846)	0.445 (0.801)	$H_T^* = 0.067$ (0.796)	4.726 (0.317)	11.131** (0.025)
$k = 3$	0.312 (0.956)	0.810 (0.847)	0.527 (0.913)	$H^* = 2.74$ (0.253)	1.819 (0.769)	1.593 (0.810)

See the *Note* in Table 1 for further explanations.

sults 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 period, 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.

III. Leading Indicator Properties of M3 within an Extended Framework

1. A P-star Model of Inflation in the Euro Area

An attempt to analyze leading indicator properties of M3 within a more 'structural' framework is made in a recent paper by Gerlach and Svensson (2000) [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:[6]

---

[6] See Groeneveld (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.

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

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

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

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

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

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

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

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

Equation (3.1a) nests the Phillips curve and price gap models of inflation, and (3.1b) specifies how inflation expectations are formed. Equation (3.1c) assumes that the monetary authority's inflation objective can be described within the available sample by a deterministic exponential trend.[7] 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. Equations(3.1d) and (3.1e) are the postulated long- and short-run money demand functions. Equatiom (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:

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

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

---

[7] The formulation (3.1c) is, in fact, taken from an earlier version of GS.

inflation objective is 1.5 percent in 1998:Q4. Focusing on the equation for π, the following estimates (standard errors in parentheses) resulted from (iii):

$$\pi_t = 0.415 \; \hat{\pi}_t - \underset{(0.089)}{0.585} \; \pi_{t-1} + \underset{(0.074)}{0.163} \; (\tilde{m} - \tilde{m}^*)_{t-1}$$
$$\underset{(-)}{}$$
$$+ \underset{(0.094)}{0.021} \; (y - y^*)_{t-1} + \hat{u}_t, \tag{3.3}$$

which can trivially be expressed as:

$$\pi_t - \hat{\pi}_{t-1} \equiv \Delta \pi_t = - 0.415 \; (\pi_{t-1} - \hat{\pi}_t)$$
$$+ 0.163 \; (\tilde{m} - \tilde{m}^*)_{t-1} + 0.021 \; (y - y^*)_{t-1} + \hat{u}_t. \tag{3.4}$$

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.[8]

2. Extending the Analysis

In this section, we build 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 II, 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.[9] As shown in

---

[8] Although the output gap is not significant at standard confidence levels in (3.3), Gerlach and Svenson (2000) 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.

[9] It may be worth noting that the real money gap can be trivially decomposed as $(\tilde{m} - \tilde{m}^*)_t \equiv (\tilde{m} - \tilde{m}^*)_{t-1} + \Delta(\tilde{m} - \tilde{m}^*)_t$. Therefore, (3.3) or (3.4) may be re-parameterized 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 GS– 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 $(\tilde{m} - \tilde{m}^*)_{t-1}$, the information content of $\Delta(\tilde{m} - \tilde{m}^*)_t$ is identical to that of $(\tilde{m} - \tilde{m}^*)_t$, 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.

Granger (1988), explicit consideration of the latter (imperfect as 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:

$$ecm1_t \equiv (m - p)_t - \underset{(0.037)}{1.158}\, y_t + \underset{(0.189)}{1.278}\, \pi_t - c$$

$$ecm2_t \equiv (R^l - R^s)_t - (R^l - R^s)^*$$

$$ecm3_t \equiv (R^s - \pi)_t - (R^s - \pi)^*\,.$$

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.[10]

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

$$\Delta \pi_t = \underset{(0.13)}{0.45} \, [(R^l - R^s) - (R^l - R^s)^*]_{t-1}$$

$$+ \underset{(0.10)}{0.54} \, [(R^s - \pi) - (R^s - \pi)^*]_{t-1} + \hat{v}_t \qquad (3.5)$$

T = 73 (1980:Q4–1998:Q4)	R^2 = 0.30	σ = 0.88% DW = 1.95
LM (1) = 1.34 (0.252)	LM(4) = 1.46 (0.231)	LM(1,4) = 0.60 (0.667)
ARCH(4) = 0.50 (0.734)	HET = 0.80 (0.553)	NORM = 0.01 (0.997)
RESET = 1.33 (0.252)	FOR(24) = 7.36 (0.999)	CHOW(24) = 0.31 (0.999)

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 ith 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 χ^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 $\Delta \pi$ and the scaled residuals. Figure 1b shows a graphical summary of results from the recursive estimation of equation (3.5) over the 1990s.

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 (ΔZ_{t-1} and $ecm1_{t-1}$) from (3.5) is not rejected either [F(8,63) = 0.88 (0.540)], 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 (0.617). 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) turn out to be relatively constant over the recent period. In these estimates due account is taken of the marked reduction in the ar-

---

[10] See section 2.1 in Coenen and Vega (2001) for a discussion of the specification of long-run money demand.

Figure 1a: *Equation 3.5: Fit and Residuals*
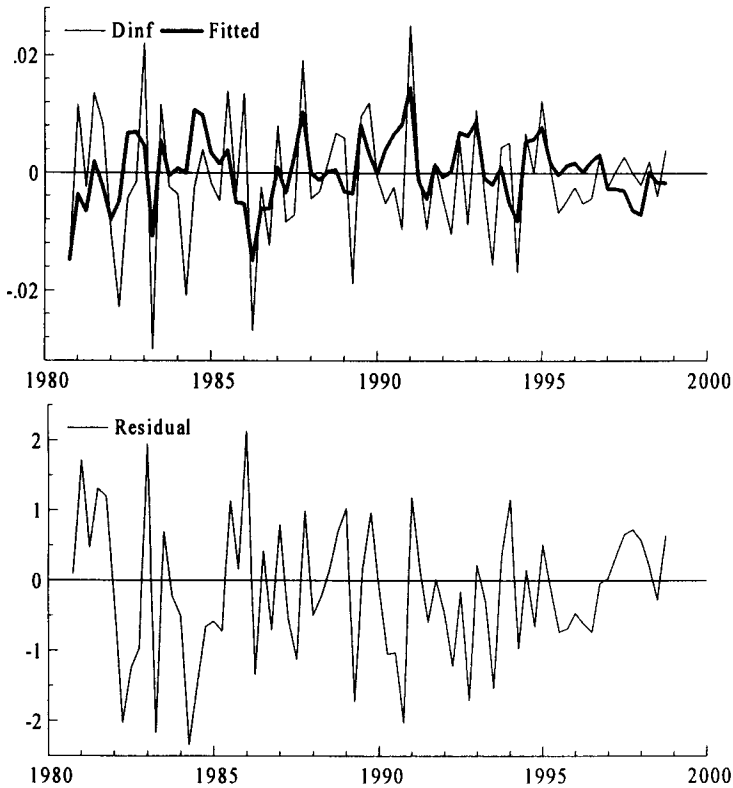
ea-wide weighted average of ex post real interest rates once exchange rate risk premia are no longer incorporated in national interest rates. This has been modeled by a step reduction in the equilibrium short-term real interest rate from 4.6 percent (approximately the weighted average of national ex post real interest rates in the period 1980:Q1 –1997:Q7) to 3.0 percent as from 1998Q1, four quarters ahead of the introduction of the single currency.[11] The equilibrium spread is estimated constant at around 0.60 percent (approximately its sample average value).

---

[11] In Gerlach and Schnabel (2000) 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 percent, with a standard error of 0.96 percent.

Figure 1b: *Equation 3.5: Recursive Estimates*
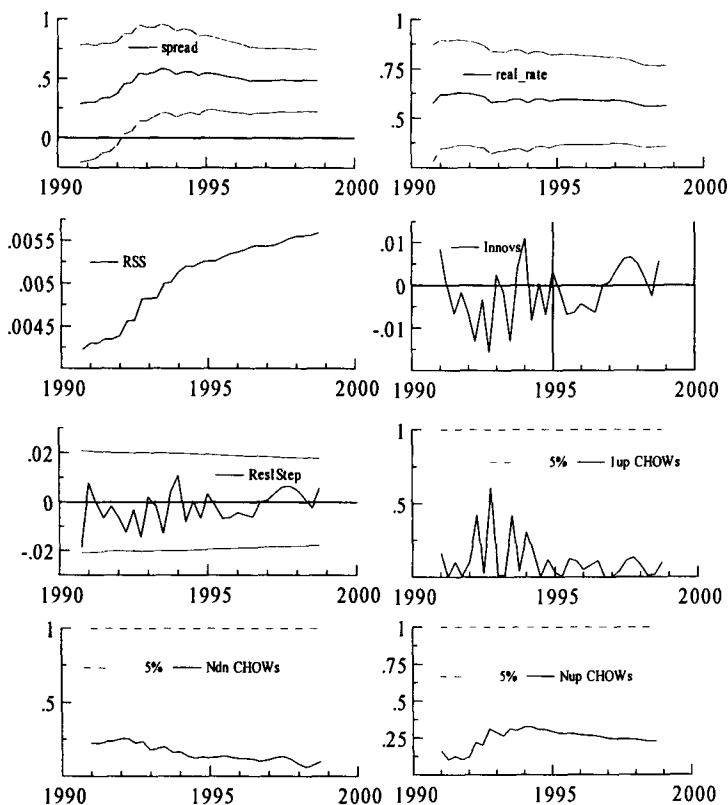
The specification in (3.5) is similar to that in Tzavalis and Wickens (1996) for the United States.[12] It shows that both the spread and the (ex post) real short-term interest rate help predict future GDP inflation in the euro area, though the low R^2 suggests that little inflation variability is indeed 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-parameterized in terms of the long-term real rate alone, indi-

---

[12] Tzavalis and Wickens (1996) review the theoretical basis for using the term spread to forecast inflation. They show for the United States 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.

cating in turn that the most of the information about future inflation is contained in the level of 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 π, the general specification encompasses specifications (3.4) and (3.5), and can be written (omitting the terms ΔZ_{t-1} for the sake of simplicity) as follows:

$$\Delta \pi_t = - \theta_0 (\pi_{t-1} - \hat{\pi}_t) + \theta_1 (\tilde{m} - \tilde{m}^*)_{t-1} + \theta_2 (y - y^*)_{t-1}$$
$$+ \theta_3 [(i^l - i^s) - (i^l - i^s)^*]_{t-1}$$
$$+ \theta_4 [(i^s_{t-1} - \hat{\pi}_t) - (i^s - \pi)^*] + \hat{v}_t \qquad (3.6)$$

where the real money gap is defined as:

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

with $k^* = k_0 + k_i (R^l - R^s)^*$, estimated using the average sample value.

Using the exponential trend postulated in GS for $\hat{\pi}$ [13] and calculating potential output by means of a Hodrick-Prescott[14] 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):

---

[13] This is here obtained by separate OLS estimation of (3.1c). It must be noticed that this procedure does not guarantee that proper account is taken of measurement errors due to the unobservable nature of $\hat{\pi}$, which could have been tackled by further extending the system to include (3.1c). However, the system estimates provide the implausible result that the implicit inflation objective is 0.1 percent 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 equal 1.5 percent is imposed, little appears to be gained from estimating (3.1c) jointly with the rest of equations in the system.

[14] 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 (1999) in the context of monetary policy rules) although it is certainly not exempt from criticism (see, for instance, King and Rebelo (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.

$$\Delta\pi_t = -\underset{(0.106)}{0.783}\,(\pi_{t-1} - \hat{\pi}_t) + \underset{(0.072)}{0.196}\,(\tilde{m} - \tilde{m}^*)_{t-1}$$

$$+ \underset{(0.130)}{0.262}\,(y - y^*)_{t-1} + \underset{(0.073)}{0.249}\,[(i^s_{t-1} - \hat{\pi}_t) - (i^s - \pi)^*] + \hat{v}_t \qquad (3.7)$$

|		
T = 73 (1980:Q4 – 1998:Q4)	$R^2 = 0.45$	$\sigma = 0.79\%$ DW = 1.92
LM(1) = 0.72 (0.399)	LM(4) = 0.85 (0.359)	LM(1,4) = 0.68 (0.607)
ARCH(4) = 0.14 (0.968)	HET = 0.221 (0.986)	NORM = 1.23 (0.541)
RESET = 0.15 (0.904)	FOR(24) = 8.5 (0.999)	CHOW(24) = 0.31 (0.998)

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 (ΔZ_{t-1} and $ecm2_{t-1}$) from (3.7) is not rejected either [F(6,63) = 1.31 (0.315)]. 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 $\Delta\pi$ and the scaled residuals. Figure 2b shows a graphical summary of results from the recursive estimation of the equation over the 1990s.

   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 $(\tilde{m} - \tilde{m}^*)_{t-1}$ 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 (3.1a + 3.1b). Yet, the presence of additional terms in (3.7) indicates that the extended P-star model given by (3.1a + 3.1b) is badly mis-specified, though estimates do not provide any indication on whether the source of this mis-specification 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 stylized facts encountered is far above the modest objectives set for this paper and will need to be addressed in future research.

3. What Leading Information in the Real Money Gap?

In this section we further investigate the main conclusion in Section III.2 (namely that the real money gap has substantive predictive power

Figure 2a: *Equation 3.7: Fit and Residuals*
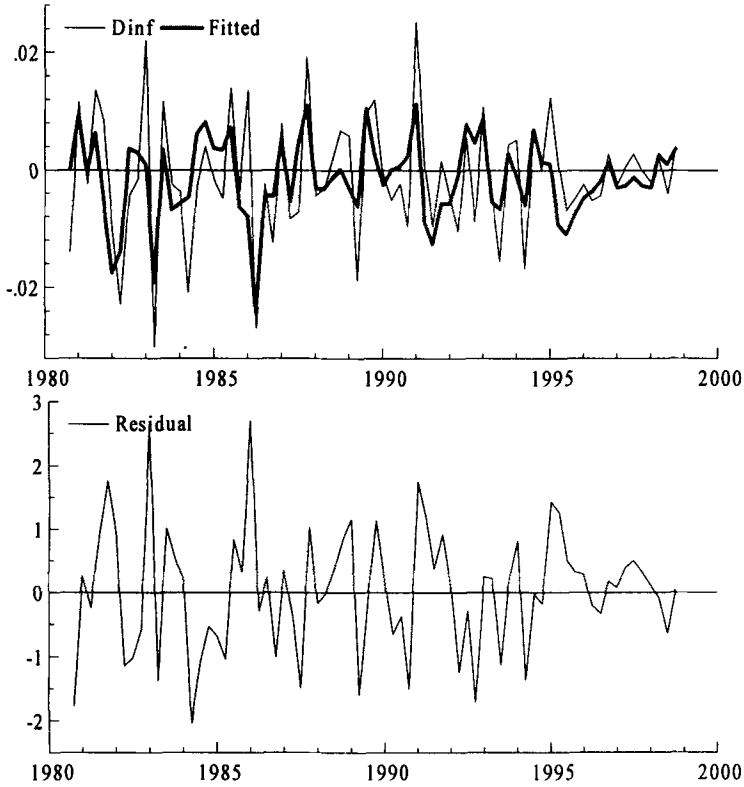
for future inflation in the euro area) by analyzing the correlations at different time horizons between $(\pi - \hat{\pi})_t$ and the various indicators included in (3.7). Results in this respect are reported in Table 3. More precisely, the table shows estimates of (3.8) below for forecasting horizons 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.[15]

---

[15] It should be noted that for $h=1$, equation (3.8) is just a trivial re-parameterization 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).

Figure 2b: *Equation 3.7: Recursive Estimates*
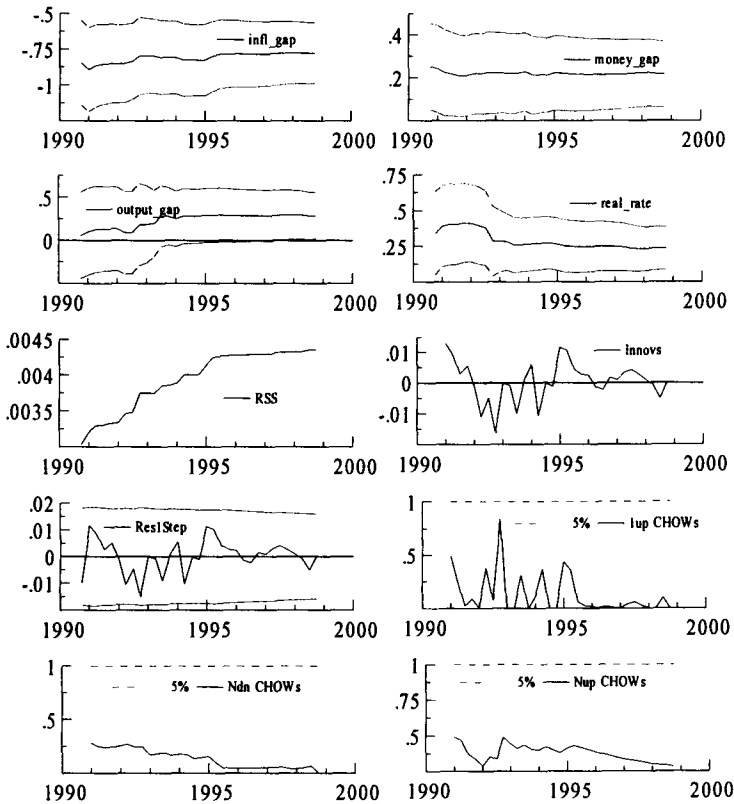
The results in Table 3 highlight the significant positive association that exists between the real money gap and future inflation. This significant correlation appears to extend over five to six quarters and reaches a maximum at the three-to-four quarter horizon. Consistent with the findings in Section III.2, the real money gap does not come out of the analysis in the table as a sufficient statistic for future inflation. The two other 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 lead-

Table 3: *Forecasting Equations for Inflation, Estimated Coefficients at Different Forecasting Horizons*

	Horizon											
	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=6$	$h=7$	$h=8$	$h=9$	$h=10$	$h=11$	$h=12$
α_1^h	0.217**	-0.012	-0.124	-0.053	-0.147	-0.159	-0.114	-0.127	-0.146	-0.172	-0.191	-0.035
α_2^h	0.196*	0.283*	0.355**	0.341**	0.316**	0.205*	0.028	-0.119	-0.076	-0.076	-0.087	-0.206
α_3^h	0.262**	0.412**	0.533**	0.588**	0.652**	0.524**	0.234	0.096	0.178	0.240	0.211	0.161
α_4^h	0.249*	0.255*	0.196**	0.097	0.113	0.146	0.211	0.141	0.212**	0.149	0.096	-0.003
$\bar{R}^2$	0.52	0.45	0.41	0.33	0.32	0.21	0.11	0.11	0.09	0.06	0.03	0.07

Note: The table shows estimates of equation (3.8) in the text for forecasting horizons ranging from $h=1$ to $h=8$ quarters: $(\pi - \hat{\pi})_{t+h-1} = \alpha_1^h(\pi_{t-1} - \hat{\pi}_t) + \alpha_2^h(\tilde{m} - \tilde{m}^*)_{t-1} + \alpha_3^h(y - y^*)_{t-1} + \alpha_4^h[(i_{t-1}^s - \hat{\pi}_t) - (i^s - \pi)^*] + \nu_{t+h-1}$, with ν_{t+h-1} following a MA$(h-1)$ process. ** and * indicate significance at 5 and 10 percent confidence levels, respectively.

ing 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 analyzed 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.

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 from the "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 sensitive results are to changes in our empirical measure of the un-observables.

With a view to analyzing 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 et al. (2001)[16] and several measures obtained by smoothing real GDP through the application of a standard Hodrick-Prescott (HP) filter with parameter $\lambda = 1600, 5000$ and $15\,000$, 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 $\hat{\pi}$.

---

[16] Potential output is estimated therein from a Cobb-Douglas production function with smoothed Solow residuals.

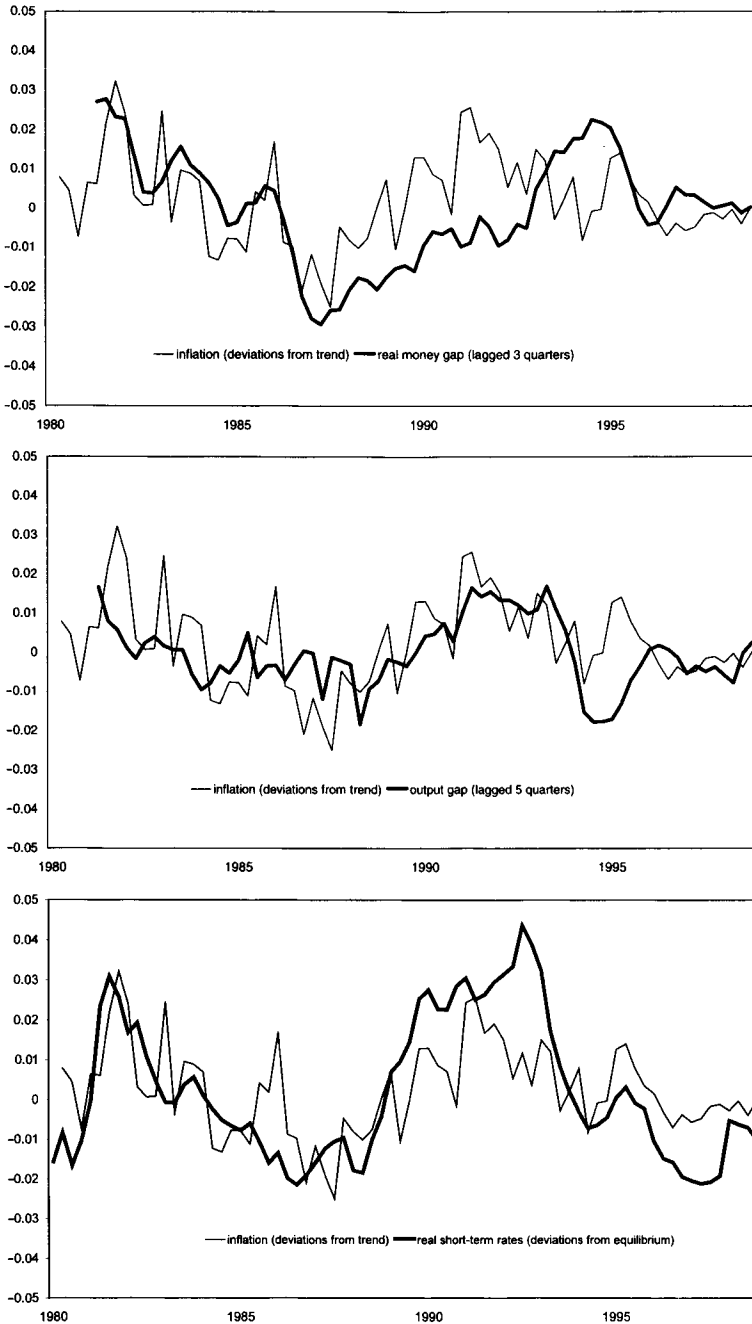Figure 3: *Inflation Deviations from "Target" and Various Gap Indicators*

Table 4: *Forecasting Equation for Inflation, Estimated Coefficients at Different Forecasting Horizons*

| | π = deterministic exponential trend | | | | y^* = HP (1600) | | | |
	y^* = AWM	y^* = HP (1600)	y^* = HP (5000)	y^* = HP (15 000)	π = determ. exp. trend	$\hat{\pi}$ = HP (1600)	$\hat{\pi}$ = (5000)	$\hat{\pi}$ = HP (15 000)
θ_0	−0.774** (0.110)	−0.812** (0.108)	−0.811** (0.107)	−0.809** (0.107)	−0.812** (0.108)	−0.831** (0.114)	−0.827** (0.112)	−0.814** (0.110)
θ_1	0.148** (0.068)	0.210** (0.076)	0.197** (0.070)	0.170** (0.065)	0.210** (0.076)	0.064 (0.056)	0.109* (0.062)	0.175** (0.073)
θ_2	0.065 (0.086)	0.264* (0.134)	0.154 (0.096)	0.095 (0.083)	0.264* (0.134)	0.195 (0.132)	0.245* (0.136)	0.291** (0.141)
θ_3	0.340** (0.141)	0.211 (0.133)	0.291** (0.132)	0.320** (0.134)	0.211 (0.133)	0.226 (0.136)	0.210 (0.136)	0.188 (0.138)
θ_4	0.446** (0.092)	0.348** (0.103)	0.362** (0.100)	0.366** (0.101)	0.348** (0.103)	0.271** (0.116)	0.282** (0.113)	0.281** (0.113)
σ (%)	0.80	0.79	0.78	0.78	0.79	0.80	0.80	0.80
R^2	0.45	0.47	0.48	0.48	0.47	0.45	0.46	0.46

Note: AWM stands for the measure of potential output included in the ECB Area-Wide Model. HP stands for Hodrick-Prescott filter, for various smoothing parameters λ. ** and * indicate significance at 5 and 10 percent confidence levels, respectively.

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 above. 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 (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.

IV. Encompassing Rival Models

In Section III.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. Since "true" models may be expected to be congruent, congruency is an important operational criterion that may help discriminate between empirical models. However, a model which is congruent is not necessarily 'true'[17] 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.

---

[17] 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 III.2.

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, 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.[18]

In the present context, one example of a non-nested congruent model for GDP inflation can be found in Fagan et al. (2001). 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 labor costs, changes in import prices and deviations of real trend unit labor costs from equilibrium. Estimation of the model over the sample 1980:Q4–1997:Q4 provides the following results:

$$\Delta\pi_t = 0.021 - 0.776\ \pi_{t-1} + 0.140\ \Delta w_t^* + 0.068\ \Delta w_{t-1}^*$$
$$\underset{(0.005)}{} \quad \underset{(0.109)}{} \quad \underset{(0.043)}{} \quad \underset{(0.047)}{}$$

$$+ 0.136\ \Delta w_{t-2}^* + 0.024\ \Delta pm_{t-1} - 0.274\ (p - w^* - \theta)_{t-1}$$
$$\underset{(0.045)}{} \quad \underset{(0.014)}{} \quad \underset{(0.076)}{} \tag{4.1}$$

T = 69 (1980:Q4–1997:Q4)	$R^2 = 0.52$	$\sigma = 0.78\%$ DW = 2.02
LM(1) = 0.22 (0.641)	LM(4) = 0.23 (0.634)	LM(1,4) = 0.54 (0.709)
ARCH(4) = 0.48 (0.750)	HET = 0.38 (0.964)	NORM = 4.12 (0.127)
RESET = 1.55 (0.218)	FOR(24) = 16.019 (0.887)	CHOW(24, 37) = 0.535 (0.945)

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

With a view to discriminating between the two contending models, Mizon and Richard's (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 (0.179) when the null is (3.7) and F(4,58) = 0.599 (0.665) when the null is (4.1). Therefore, based on within-sample information, there appears to be little to choose between both models[19]:

---

[18] 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).

[19] Note also that the fits of the two models – as measured by their respective residual standard errors – are close.

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:Q1 – 1997:Q4.

We also tested for out-of-sample forecast encompassing, with forecasts under models (3.7) and (4.1) generated by 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[20]:

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

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

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 reports the corresponding root mean square forecast errors at the one- to eight-quarter forecast horizons together with some tests for forecasting encompassing.[21] When compared to the equation

---

[20] 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 (2002) in relation to the measurement of the output gap.

[21] See Chong and Hendry (1986) and Harvey et al. (1998).

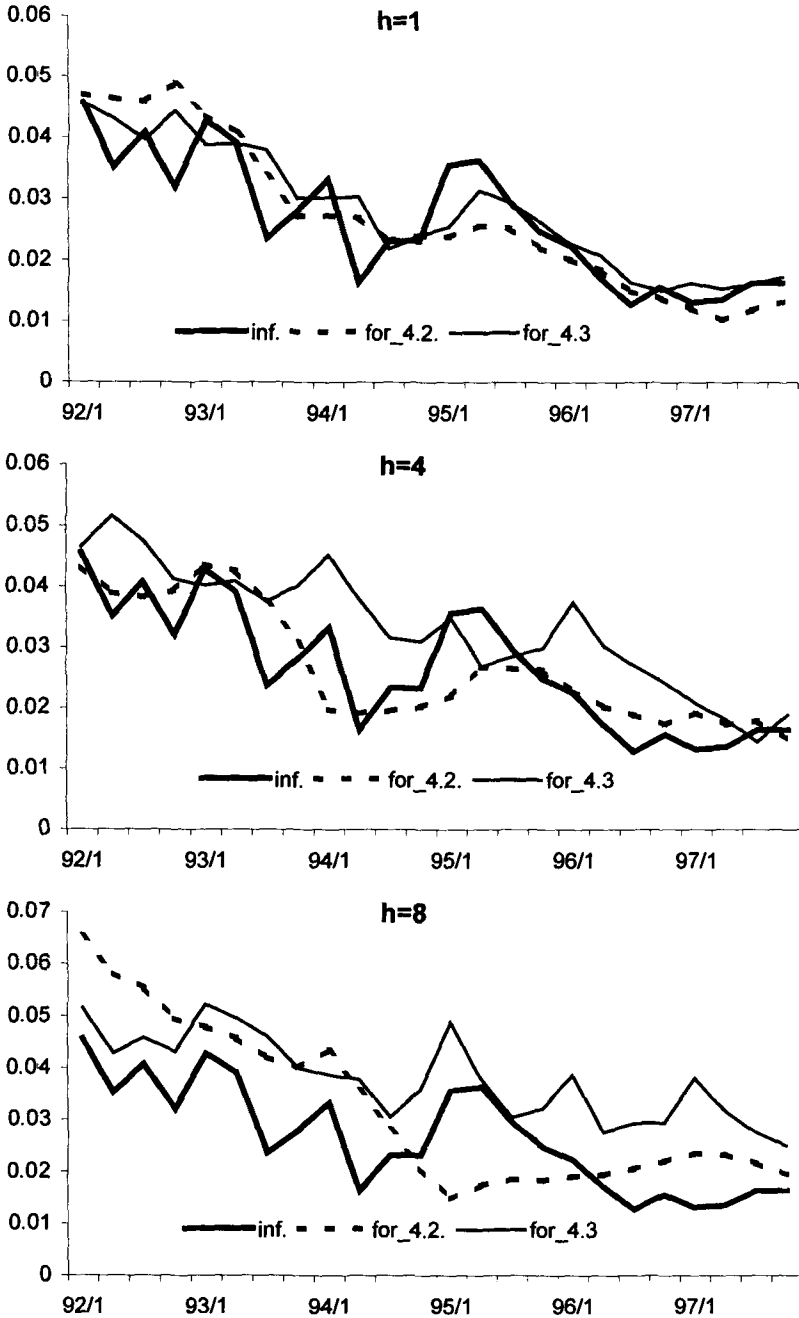Figure 4: *Inflation Forecasts at Different Horizons*

Table 5: *Root Mean Square Forecast Errors at the One-*
to-Eight-Quarter Forecast Horizons (h) and Tests
for Forecasting Encompassing

h	α	Model 4.2		Model 4.3	
		RMSE	H_0: $\alpha = 0$	RMSE	H_1: $\alpha = 1$
1	0.830 (0.348)	0.653	0.0171**	0.585	0.6251
2	0.575 (0.176)	0.656	0.0011**	0.623	0.0156**
3	0.241 (0.053)	0.638	0.0000**	0.834	0.0000**
4	0.178 (0.098)	0.614	0.0701*	0.992	0.0000**
5	0.335 (0.121)	0.827	0.0058**	1.033	0.0000**
6	0.454 (0.094)	1.094	0.0000**	1.148	0.0000**
7	0.447 (0.157)	1.145	0.0043**	1.208	0.0004**
8	0.472 (0.198)	1.259	0.0173**	1.293	0.0078**

Note: ** and * indicate significance at 5 and 10 percent confidence levels.

standard errors reported above, the results for $h = 1$ imply that both models fare particularly well during the period under analysis 1992: Q1 – 1997:Q4. In any event, Chow test statistics do not detect any structural break: Chow (24,41) = 0.508 (0.960) in the case of (3.7) and Chow (24,37) = 0.535 (0.945) in the case of (4.1). As regards forecasting encompassing, the table shows estimates of α (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:

$$\varepsilon_{1t}^h = \alpha (\varepsilon_{1t}^h - \varepsilon_{2t}^h) + \varsigma_t, \tag{4.4}$$

where ε_1^h and ε_2^h ($h = 1, \dots, 8$) stand for the corresponding forecast errors obtained from (4.2) and (4.3), respectively. Rejection of one of the null hypotheses provides evidence in favor 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 mis-specified.

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.[22] Indeed, for $h = 1$ model (4.2) appears to incorporate little information which is not already contained in model (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.

V. 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 II the leading indicator properties of broad money M3 were analyzed by conducting tests that money does not Granger-cause prices in the context of a cointegrated VAR system that has recently been proposed to investigate 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 the information set employed therein. 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.

We further investigated in Section III the leading indicator properties of broad money M3 by looking at the recent claim that – in the context of a P-star model – 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

---

[22] 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.

this empirical finding can be reproduced when the information set employed in Section II is augmented to include variables such as 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, we believe 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 stylized fact appears to be quite, though not totally, robust to the use of a number of empirical measures for the two unobservables.

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 III, 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 P-star model of inflation. This in turn suggests that the treatment of inflation expectations and most possibly the measurement of the monetary policy authorities' implicit inflation objective needs at least to be refined.

Finally, in Section IV we compared the model developed in the previous section with an alternative model for GDP inflation in the euro area with no explicit role assigned to money. In the latter, 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 labor costs, changes in import prices and deviations of real trend unit labor costs 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 the rival model) is, however, rejected at all horizons other than $h = 1$. On the basis of this evidence, we conclude that each model appears to have strengths 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

Chong, Y. Y., and D. F. Hendry (1986). Econometric Evaluation of Linear Macro-Economic Models. *Review of Economic Studies* 53 (4): 671–690.

Clements, M. P., and D. F. Hendry (1996). Multi-Step Estimation for Forecasting. *Oxford Bulletin of Economics and Statistics* 58 (4): 657–684.

Coenen, G., and J. L. Vega (2001). The Demand for M3 in the Euro Area. *Journal of Applied Econometrics* 16 (6): 727–748.

Dolado, J. J., and H. Lütkepohl (1996). Making Wald Tests Work in Cointegrated VAR Systems. *Econometric Reviews* 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 F. S. Mishkin (1997). Is There a Role for Monetary Aggregates in the Conduct of Monetary Policy? *Journal of Monetary Economics* 40 (2): 279–304.

European Central Bank (1999). The Stability-Oriented Monetary Policy Strategy of the Eurosystem. *ECB Monthly Bulletin* (January): 39–50.

Fagan, G., J. Henry, and R. Mestre (2001). An Area-Wide Model (AWM) for the Euro Area. European Central Bank, Working Paper 42. Frankfurt/M.

Gerlach, S., and G. Schnabel (2000). The Taylor Rule and Interest Rates in the EMU Area: A Note. *Economics Letters* 67 (2): 165–171.

Gerlach, S., and L. E. O. Svensson (2000). Money and Inflation in the Euro Area: A Case for Monetary Indicators? BIS Working Paper 98, Basle (also published as NBER Working Paper 8255, Cambridge, Mass.).

Granger, C. W. J. (1988). Causality, Cointegration and Control. *Journal of Economics, Dynamic and Control* 12 (2–3): 551–558.

Groeneveld, J. M. (1998). *Inflation Patterns and Monetary Policy. Lessons for the European Central Bank.* Cheltemham: Edward Elgar.

Harvey, A. C., and A. Jaeger (1993). Detrending, Stylised Facts and the Business Cycle. *Journal of Applied Econometrics* 8 (3): 231–247.

Harvey, D. I., S. J. Leybourne, and P. Newbold (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 B. Cornet and H. Tulkens (eds), *Contributions to Operations Research and Econometrics. The XX Anniversary of CORE.* Cambridge, Mass.: MIT Press.

Issing, O., V. Gaspar, I. Angeloni, and O. Tristani (2001). *Monetary Policy in the Euro Area: Strategy and Decision-Making at the European Central Bank.* Cambridge, UK.: Cambridge University Press.

King, R. G., and S. Rebelo (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* 105 (3): 815–828.

Mizon, G. E., and J. F.Richards (1986). The Encompassing Principle and Its Application to Testing Non-Nested Hypotheses. *Econometrica* 54 (3): 657–678.

Orphanides, A., and S. van Norden (2002). The Unreliability of Output Gap Estimates in Real Time. Board of Governors of the Federal Reserve System. *The Review of Economics and Statistics*, forthcoming.

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

Sims, C. A. (1972). Money, Income and Causality. *American Economic Review* 62 (4): 540–552.

Stock, J. H., and M. W. Watson (1999). Forecasting Inflation. *Journal of Monetary Economics* 44 (2): 293–335.

Taylor, J. B. (1999). A Historical Analysis of Monetary Policy Rules. In J. B. Taylor (ed.), (1999). *Monetary Policy Rules*. Chicago: The University of Chicago Press.

Toda, H. Y., and P. C. B. Phillips (1993). Vector Autoregression and Causality. *Econometrica* 63 (4) 1367–1393.

Toda, H. Y., and P. C. B. Phillips (1994a). Vector Autoregression and Causality: A Theoretical Overview and Simulation Study. *Econometric Reviews* 13 (2): 259–285.

Toda, H. Y., and T. Yamamoto (1994). Statistical Inference in Vector Autoregressions with Possibly Integrated Processes. *Journal of Econometrics* 66 (1–2): 225–250.

Tzavalis, E., and M. R. Wickens (1996). Forecasting Inflation from the Term Structure. *Journal of Empirical Finance* 3 (1): 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*. Chicago: University of Chicago Press for NBER.

Yamada, H., and H. Y. Toda (1998). Inference in Possibly Integrated Vector Autoregressive Models: Some Finite Sample Evidence. *Journal of Econometrics* 86 (1): 55–95.

* * *

Abstract: The Information Content of M3 for Future Inflation in the Euro Area. – The information content of M3 for future inflation in the euro area is investigated from a number of perspectives. Our results confirm that a significant positive association exists between the real money gap and future inflation up to five to six quarters ahead. It is also shown that, although the extended P-star model outperforms the rival model in some respect, the hypothesis that no useful information is contained in rival evidence can be rejected. JEL no. C32, C50, E30, E40