Economics Discussion Papers
The Open-Access, Open-Assessment E-Journal

Discussion Paper 2008-15
April 10, 2008

# Evaluating the New Keynesian Phillips Curve under VAR-Based Learning

*Luca Fanelli*
*University of Bologna*

**Please cite the corresponding journal article:**
http://www.economics-ejournal.org/economics/journalarticles/2008-33

**Abstract:**

This paper proposes the econometric evaluation of the New Keynesian Phillips Curve (NKPC) in the euro area, under a particular specification of the adaptive learning hypothesis. The key assumption is that agents' perceived law of motion is a Vector Autoregressive (VAR) model, whose coefficients are updated by maximum likelihood estimation, as the information set increases over time. Each time new data is available, likelihood ratio tests for the crossequation restrictions that the NKPC imposes on the VAR are computed and compared with a proper set of critical values which take the sequential nature of the test into account. The analysis is developed by focusing on the case where the variables entering the NKPC can be approximated as nonstationary cointegrated processes, assuming that the agents' recursive estimation algorithm involves only the parameters associated with the short run transient dynamics of the system. Results on quarterly data relative to the period 1981–2006 show that: (i) the euro area inflation rate and the wage share are cointegrated; (ii) the cointegrated version of the 'hybrid' NKPC is sharply rejected under the rational expectations hypothesis; (iii) the model is supported by the data over relevant fractions of the chosen monitoring period, 1986–2006, under the adaptive learning hypothesis, although this evidence does not appear compelling.

Paper submitted to the special issue "Using Econometrics for Assessing Economic Models" edited by Katarina Juselius.

*JEL: C32, C52, D83, E10*
*Keywords: Adaptive learning; cointegration; cross-equation restrictions; forward-looking model of inflation dynamics; New Keynesian Phillips Curve; Recursive Least Squares; VAR; VEqC*

*Correspondence: Luca Fanelli, Department of Statistical Sciences, University of Bologna, via Belle Arti 41, I–40126 Bologna, Italy. email: luca.fanelli@unibo.it, ph: +39 0541434303, fax: +39 051 232153.*

*I wish to thank Katarina Juselius, Barbara Rossi and Davide Delle Monache for helpful comments and suggestions on earlier drafts of the paper.*

www.economics-ejournal.org/economics/discussionpapers

# 1 Introduction

There is growing awareness among applied researchers that the New Keynesian Phillips Curve (NKPC),[1] which plays a dominating role in the monetary policy literature, provides a poor explanation of inflation dynamics and persistence in developed countries, see Fuhrer and Moore (1995), Fuhrer (1997), Ruud and Whelan (2005*a*, 2005*b*, 2006), Boug et al. (2007) and Fanelli (2008), just to mention a few. Many explanations have been provided, including identification issues (Mavroeidis, 2005; Nason and Smith, 2005; Boug et al., 2007), dynamic misspecification (Bårdsen et al. 2004; Dees et al., 2008; Fanelli, 2008*a*) and neglected nonstationarity (Juselius, 2006; Fanelli, 2008*a*), however, the poor empirical performance of models featuring the rational expectations hypothesis (REH), is pervasive in macroeconomics and finance.

The NKPC is grounded on the REH, and its empirical investigation usually maintains that the REH holds. Aside from the difficulties of disentangling forward- from backward-looking behaviour on empirical grounds (Hendry, 1988), many authors argue that in practice agents depart from the REH, and display either imperfect knowledge (Goldberg and Frydman, 2007) or 'bounded' rationality, see Pesaran (1987), Sargent (1999) and Evans and Honkapohja (1999, 2001). For instance, focusing on exchange rate markets, Goldberg and Frydman (1996) argue that an alternative, and arguably more plausible, assumption is that economic agents have only imperfect knowledge of the true relationship between exchange rates and fundamentals, and appeal mainly to qualitative rather than quantitative knowledge about the economy. In the monetary policy framework, the idea that inflation expectations may not be rational and that deviations from the REH may represent a remarkable source of inflation persistence is well a recognized fact, see Roberts (1997) and Milani (2005). Sargent (1999), Orphanides and Williams (2005) and Primicieri (2006) are examples where rational policy-makers learn about the behavior of the economy in real time, and set stabilization policies conditional on their current beliefs.[2]

The traditional approach to modelling boundedly rational expectations assumes that agents behave as econometricians when making forecasts and use adaptive learning algorithms to update their beliefs (Evans and Honkapohja 1999, 2001). This means that they estimate and update the parameters of their forecasting model - the perceived law of motion - according to recursive rules.[3] Replacing expectations in the forward-looking model, with the forecasts implied by

[1] All acronyms used in the paper are reported in Table 1.

[2] Moreover, optimal policy rules designed under the REH may not perform satisfactorily if instead of having rational expectations private agents follow learning rules: Bullard and Mitra (2002) show that in this case the stability of the Taylor-type rules can not be taken for granted, see also Evans and McGough (2005). Evans and Honkapohja (2003*a*, 2003*b*) show how optimal monetary can be designed in these situations, provided that the mechanism generating private expectations is suitably incorporated into the model.

[3] In this context learning is 'adaptive' rather than 'optimal', because it ignores the feedback from the learning

the perceived law of motion, yields the so-called actual law of motion, which reads as the agents' data generating process. Under certain conditions, it has been found that expectations in these models can converge to a rational expectations equilibrium (or to a restricted perception equilibrium, Evans and Honkapohja, 2001), that means that in the limit the actual law of motion is indistinguishable from the model solution obtained under the REH.

The notion of adaptive learning has been applied in the macroeconomic literature mainly as a selection criterion when the rational expectations model generates multiple solutions, and in connection with the concept of 'learnability' and stability of rational expectations equilibria. On the econometric side, however, although the existing literature typically focuses on the problem of estimating the actual law of motion through recursive (possibly Bayesian) methods, little is known about inference.

The focus of this paper is not on the agents' estimation problem, but rather on the problem of testing a forward-looking model, as the NKPC, under the adaptive learning hypothesis (ALH).[4] So far, the issue of testing the data adequacy of models based on particular specifications of the ALH has been disregarded. There are at least two related reasons for changing this state of the art. First, if one finds that a given forward-looking model is rejected by the data under the REH, and is supported under the ALH, it can be reasonably argued that adaptive learning represents a more sensible description of the way agents form their forecasts in the economy; in this respect, the ALH can reconcile a class of forward-looking models with the data. Second, testing the econometric implications of a given forward-looking model under the ALH, allows the researcher to assess whether the claimed process of convergence to a potential rational expectation equilibrium (or restricted perception equilibrium), is consistent with the observed data or not.

To our knowledge Fanelli and Palomba (2007) and Fanelli (2008*b*) are the only existing contributions in which the NKPC has been investigated by Vector Autoregressive (VAR) models and recursive likelihood-based methods under a particular specification of the ALH.[5] In this paper we generalize and extend those contributions by focusing on the case in which the inflation rate and/or its driving variable(s) can be approximated as nonstationary cointegrated processes.

---

rule on the actual low of motion. In this paper we focus on the issue of learning from a purely econometric point of view.

[4] To our knowledge, In-Koo and Kasa (2005) is a contribution where the issue of learning and model validation is explicitly addressed by taking the agents' point of view.

[5] Although the literature provides several examples where dynamic stochastic general equilibrium models comprising NKPC-like supply equations are investigated under adaptive learning, to our knowledge, Milani (2005) is the only contribution where 'the econometric analysis of the NKPC under learning rules' is explicitly addressed. However, the analysis in Milani (2005) is based on the estimation of the resulting actual law of motion, and not on testing the data adequacy of the model under the ALH. More on this in Section 4.

The investigation of the NKPC is carried out under the following set of assumptions: **(i)** agents use VAR models including inflation and its driving variables as their perceived law of motion; **(ii)** VAR coefficients are updated recursively through the reiterate application of maximum likelihood estimation; **(iii)** when the variables entering the NKPC are nonstationary and cointegrated, the agents' learning rule involves only the short run adjustment coefficients of the Vector Equilibrium Correction (VEqC) counterpart of the VAR, and not the cointegration parameters.

We discuss in detail each of the three assumptions and related implications.

**Assumption (i)**: the idea that agents use VARs to form their expectations is not new in the literature, and is known as the 'VAR expectations' hypothesis (hereafter VEH), see Brayton et al. (1997), Johansen and Swensen (1999), Kozicki and Tinsley (1999), Branch (2004). Under the VEH, it is possible to derive a set of cross-equation restrictions between the VAR coefficients and the NKPC along the lines of Fanelli (2008*a*). The VEH coincides with the REH when the (determinate, if any) solution of the rational expectations model is nested within the VAR serving as the statistical model for the data. In general, however, we *do not assume* that the perceived law of motion *necessarily coincides* with the minimum state variable solution (McCallum, 1983, 2003) of the system.

**Assumption (ii)**: given the recursive nature of the estimation problem implied by the ALH, by construction the cross-equation restrictions between the VAR coefficients and the NKPC have a sequential nature. These restrictions can be recursively tested through likelihood-ratio tests obtained by estimating the VAR both unrestrictedly and subject to the restrictions, over the monitoring period. From the inferential point of view, however, standard asymptotic theory can not be applied due to the law of iterated logarithms (Inoue and Rossi, 2005), hence the critical values must take the sequential nature of the test into account.

**Assumption (iii)**: when the variables entering the NKPC can be approximated as nonstationary cointegrated processes, the evaluation of the model under the ALH is pursued by resorting to suitable transformations of the VEqC counterpart of the VAR. The analysis is developed under the hypothesis that the agents' learning rule involves only the parameters associated with the short run transient dynamics of the system, and not the cointegration parameters. This assumption, which can be relaxed in future research, is consistent with the idea that since the cointegration parameters are related to the long run (steady state) solution of the model, in principle they are not learnable during the transition process to equilibrium.

The proposed method, based on the three assumptions above, is applied to investigate the NKPC in the euro area, using quarterly data for the period 1981-2006. We proxy firms real marginal costs by the wage share as in Galí et al. (2001), and include also a short term nominal interest rate in the system, given the influence of interest rates on marginal costs through the so-called cost channel of monetary transmission (Chowdhury et al., 2005). In previous research, Bårdsen et al. (2004), O'Reilly and Whelan (2005) and Fanelli (2008*a*) have shown that the 'hybrid' formulation of the NKPC under the REH does not capture inflation dynamics successfully in the euro area. In this paper we find that the nonstationarity of variables is an issue, and that the inflation rate and the wage share are cointegrated over the period 1981-2006. Moreover, we find that the NKPC is sharply rejected under the VEH, but receives some empirical support over a large fraction of the chosen monitoring period, 1986-2006, when the model is recursively tested under the ALH; this evidence, however, is not compelling. We discuss the implications of this result, and compare it with the findings of other authors.

The rest of the paper is organized as follows. Section 2 introduces the NKPC and discusses two possible representations of the model when variables can be approximated as nonstationary processes. Section 3, which is divided into two subsections, proposes the econometric investigation of the cointegrated NKPC under the ALH. Section 4 applies the method to investigate the NKPC in the euro area data and discusses the relative implications. Section 5 contains some concluding remarks.

| Mnemonic | Definition |
|---|---|
| ALH | Adaptive Learning Hypothesis |
| AWM | Area Wide Model |
| BFGS | Broyden, Fletcher, Goldfarb, Shanno method, see Fletcher (1987) |
| MDS | Martingale Difference Sequence |
| NKPC | New Keynesian Phillips Curve |
| REH | Rational Expectations Hypothesis |
| VAR | Vector Autoregressive |
| VEH | Vector Expectations Hypothesis |
| VEqC | Vector Equilibrium Correction |

Table 1: Acronyms used in the paper.

## 2 Model and cointegration implications

A variety of pricing environments within the New Keynesian tradition give rise to the following 'hybrid' version of the NKPC

$$\pi_t = \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \lambda x_t + u_t \tag{1}$$

where $\pi_t$ is the inflation rate at time $t$, $x_t$ is a scalar explanatory variable related to firms' real marginal costs, and $E_t\pi_{t+1}$ indicates the expected value of $\pi_{t+1}$ formed at time $t$ on the basis of the available information summarized in the (non decreasing) sigma-field $\Omega_t \subseteq \Omega_{t+1}$, i.e. $E_t\pi_{t+1} = E(\pi_{t+1} \mid \Omega_t)$. $\gamma_f$, $\gamma_b$ and $\lambda$ are the structural parameters.

The properties of the process generating $x_t$ in (1) are crucial for the derivation of the model solution and for the identification of the parameters; for ease of exposition, we leave first the process generating $x_t$ unspecified; in the next section, the dynamics of $x_t$ will follow a VAR model for $\pi_t$ and $x_t$ (and possibly other variables). The NKPC in (1) has been specified by including a disturbance term $u_t$. In the literature, it is a common practice to assume that the shock $u_t$ obeys autoregressive dynamics (typically an AR(1) process), but here we assume that $u_t$ is martingale difference sequence (MDS) with respect to $\Omega_t$, for two reasons. First, we interpret $u_t$ as a term capturing unexplained (transitory) deviations from the theory (Kurmann, 2007). Second, modelling $u_t$ as a persistent processes without any indication from the theory would represent an *ad hoc* assumption, not derived from first principles.

The structural parameters $\gamma_f$, $\gamma_b$ and $\lambda$ of model (1) can be generally expressed as nonlinear functions of 'primitive' parameters of the model, related to consumers and firms preferences. For instance, in the Calvo model of Galí and Gertler (1999) and Galí et al. (2001) one has

$$\gamma_f = \frac{\rho\theta}{\theta + \omega[1-\theta(1-\rho)]} \tag{2}$$

$$\gamma_b = \frac{\varpi}{\theta + \varpi[1-\theta(1-\rho)]} \tag{3}$$

$$\lambda = \frac{(1-\varpi)(1-\theta)(1-\theta\rho)}{\theta + \varpi[1-\theta(1-\rho)]} \tag{4}$$

where $\rho$ is firms' discount factor, $0 < \varpi < 1$ is the fraction of backward-looking firms in the economy that change their prices following rule-of-thumb behavior, and $0 < \theta < 1$ is the probability that a firm will be unable to change its price in a given period, so that $(1-\theta)^{-1}$ is the average duration over which a price is fixed. By construction $\gamma_f + \gamma_b \leq 1$.

The empirical assessment of the NKPC (1) under the REH has attracted a great deal of research. Abstracting from 'limited-information' methods, two 'full-information' likelihood-based approaches have been applied. The former is based on the explicit specification of the process generating $x_t$, and the derivation of the (possibly unique) reduced form solution of the system. The process generating $x_t$ can be either a reduced form as in Pesaran (1987, Chap. 7), and Fuhrer and Moore (1995) and Fuhrer (1997), or a structural equation drawn from a 'small-scale' dynamic stochastic general equilibrium model of monetary policy as in Lindè (2005). Here it is of key importance to envisage whether the model solution under the REH is determinate or indeterminate, see Lubik and Schorfheide (2004).

The latter approach, based on the VEH, works under the (implicit) assumption that the data generating process belongs to the correctly specified VAR for $Z_t = (\pi_t \colon x_t)'$. The application of the method of undetermined coefficients allows to retrieve a set of cross-equation restrictions between the VAR and NKPC parameters, which can be used to estimate and test the model: the idea is that if the NKPC is the true inflation model of the economy, then the VAR should support the implied cross-equation restrictions, see Kurmann (2007) and Fanelli (2008*a*).

Aside from estimation methods, the empirical investigation of the NKPC (1) is usually carried out assuming that variables are generated by stationary processes, in line with the idea that the NKPC is derived from a dynamic stochastic general equilibrium model which is solved by log-linearizing around a steady state. We refer to Dees et al. (2008) for a comprehensive discussion of the perils of computing steady states through simple means or statistical procedures, and for possible remedies. A crucial point is that if stationarity is wrongly assumed, inference in the class of forward-looking models that can be cast in the form (1) may be misleading, see Johansen (2006) and Franchi and Juselius (2007).

Two formulations of the original NKPC model (1) are worth considering when variables are nonstationary and cointegrated. The former is based on the restriction $\gamma_f + \gamma_b = 1$, and is obtained by manipulating equation (1) in the form

$$\Delta\pi_t = \psi E_t \Delta\pi_{t+1} + \omega x_t + u_t^* \tag{5}$$

where

$$\psi = \frac{1-\gamma_b}{\gamma_b} \tag{6}$$

$$\omega = \frac{\lambda}{\gamma_b} \tag{7}$$

and $u_t^* = u_t/\gamma_b$. This representation emphasizes the role of the stationary variable $x_t$ as driving force of the inflation acceleration rate, $\Delta\pi_t$. Equation (5) fits the data when $\pi_t$ is integrated of order one (I(1)) and $x_t$ is stationary (for example the output gap should be stationary by construction).

The latter formulation is based on the assumption that $\pi_t$ and $x_t$ are both I(1) and share a common stochastic trend.[6] Provided that $\gamma_f + \gamma_b < 1$, and assuming that $\pi_t$ and $x_t$ are cointegrated with coefficients $(1, -\phi)$, the model is obtained by manipulating equation (1) in the form

$$\Delta\pi_t = \psi E_t \Delta\pi_{t+1} + \omega(\pi_t - \phi x_t) + u_t^* \tag{8}$$

[6] In Section 4 we find that the euro area inflation rate and the wage share (log of real unit labour costs) are cointegrated over the period 1981-2006. We do not attach any particular interpretation to this finding, see Fanelli (2008*a*) for details.

where

$$\psi = \frac{\gamma_f}{\gamma_b}, \tag{9}$$

$$\omega = -\frac{[1-(\gamma_f+\gamma_b)]}{\gamma_b}, \tag{10}$$

$$\phi = \frac{\lambda}{[1-(\gamma_f+\gamma_b)]}. \tag{11}$$

The interesting feature of the cointegrated NKPC (8), is that equation (11) establishes a link between the slope parameter $\lambda$, and the cointegration parameter $\phi$: in general, for given $\phi$, $\gamma_f$ and $\gamma_b$, it turns out that $\lambda = \phi[1-(\gamma_f+\gamma_b)]$ is automatically determined, see Fanelli (2008*a*).

Disentangling between the formulation (5)-(7) and the formulation (8)-(11) of the cointegrated NKPC is an empirical question, that can be addressed by investigating the cointegration properties of the system $Z_t = (\pi_t \vdots\, x_t \vdots\, a_t')'$, where $a_t$ is a $q_a \times 1$ vector of additional variables which are deemed to be relevant for the analysis. If, for instance, it is found that $Z_t$ embodies a single cointegrating relation,[7] one can test whether $\beta' Z_t$ is consistent with structure $\beta' = (0, 1, 0')$ (model (5)), or with the structure $\beta' = (1, -\phi, 0')$ (model (8)).

The two representations (8) and (5) can be nested in the more general specification

$$\Delta\pi_t^d = \psi\ E_t \Delta\pi_{t+1}^d + \omega(\beta' Z_t)^d + u_t^* \tag{12}$$

where $u_t^*$ is a MDS, and $\psi$ and $\omega$ are determined according to (6)-(7) or (9)-(10), depending on the structure of $\beta$. In (12) we have used the superscript '$d$' to remark that the variables entering the model represent deviations from constant (steady state) levels.

To investigate the NKPC (12) through 'full-information' methods, it is necessary to derive the restrictions that model (12) imposes on a VAR for $Z_t$, where the latter captures the statistical properties of the observed time series. Provided that the VAR is identifiable under the restrictions, one can estimate the model both unrestrictedly and subject to the constraints, and compute a likelihood-ratio test, see Fanelli and Palomba (2007) and Fanelli (2008*a*). In the next section we briefly review this method, and then extend it to a recursive framework to account for the implications of the ALH.

[7] Of course, when the cointegration rank is greater than one, the researcher faces the problem of identifying the additional long run relations.

# 3 The cointegrated NKPC under the adaptive learning hypothesis

For ease of exposition we divide this section into two parts. In Subsection 3.1 we review the analysis of the NKPC under the VEH, and in Subsection 3.2 we derive and discuss the implications that arise when the ALH is assumed.

## 3.1 Cross-equation restrictions under the VEH

Assume that the law of motion for the $p \times 1$ vector of observable variables, $Z_t = (\pi_t \colon x_t : a_t')'$, is given by

$$Z_t = \sum_{i=1}^{k} A_i Z_{t-i} + \Theta D_t + \varepsilon_t \tag{13}$$

where $k$ is the lag length, $Z_0$, $Z_{-1}, ...,$ $Z_{(1-k)}$ are fixed, $A_i$, $i = 1, 2, ..., k$ are $p \times p$ matrices of parameters, $D_t$ is an $d_0 \times 1$ vector of deterministic terms (constant, linear trend, deterministic dummies, etc.) with associated $p \times d_0$ matrix of parameters, $\Theta$, and $\varepsilon_t$ is a martingale difference sequence (MDS) with respect to the sigma-field $\mathcal{H}_t = \sigma(Z_t, Z_{t-1}, ..., Z_1) \subseteq \Omega_t$, with (non-singular) covariance matrix $\Sigma_\varepsilon$ and Gaussian distribution.

Given the VAR characteristic polynomial $A(L) = I_p - A_1 L - \cdots - A_k L^k$, where $L$ is the lag operator, it is assumed that the roots, $s$, of $\det[A(s)] = 0$ are such that $\mid s \mid \geq 1$, hence explosive roots are ruled out. Moreover, we maintain that when there are (unit) roots at $s = 1$, their exact number is $p - r$, where $r$, $0 < r < p$, is the cointegration rank of the system; this means that the focus is on I(1) cointegrated processes, see Johansen (1996).

The VAR (13) is here treated as the agents' forecast model (perceived law of motion), and it is assumed that the VAR lag length, $k$, fulfils the restriction $pk \geq 3 + (p - r)$. As it will be clear below, this restriction is necessary to guarantee the local identifiability of the VAR under the nonlinear cross-equation restrictions implied by the NKPC.

Given the cointegrated VAR (13), we consider the corresponding VEqC counterpart

$$\Delta Z_t = \alpha\beta' Z_{t-1} + \sum_{i=1}^{k-1} \Phi_i \Delta Z_{t-i+1} + \mu + \varepsilon_t \tag{14}$$

where $\alpha\beta' = -(I_p - \sum_{i=1}^{k} A_i)$ is the long run impact matrix with $\alpha$ and $\beta$ two $p \times r$ full rank matrices, and $\Phi_j = -\sum_{i=j+1}^{k} A_i$, $j = 1, ..., k-1$; for easy of exposition, and in line with the results of Section 4, we consider the case where $D_t = 1$ and $\Theta = \mu$, admitting the presence of deterministic linear trends in the system. Under a suitable set of identifying restrictions, the elements in the $r \times 1$ vector $\beta' Z_t$ define the cointegration relations of the system (Johansen, 1996; Juselius, 2006).

Finally, in order to derive the cross-equation restrictions between the VEqC and the representation (12) of the NKPC, it is convenient to re-arrange the variables in 'triangular' form, introducing the vector

$$W_t = \begin{pmatrix} \beta' Z_t \\ v' \Delta Z_t \end{pmatrix} \equiv \begin{pmatrix} W_{1t} \\ W_{2t} \end{pmatrix} \quad \begin{matrix} r \times 1 \\ (p-r) \times 1 \end{matrix} \tag{15}$$

where $v$ is a $p \times (p-r)$ matrix such that $\det(v'\beta_\perp) \neq 0$, and $\beta_\perp$ is the orthogonal complement of $\beta$ (Johansen, 1996). It can be proved (Paruolo, 2003, Theorem 2) that $W_t$ admits the VAR representation

$$W_t = \sum_{i=1}^{k} B_i W_{t-i} + \mu^0 + \varepsilon_t^0 \tag{16}$$

where $\varepsilon_t^0 = (\beta{,}v)'\varepsilon_t$ is a MDS with covariance matrix $\Sigma_{\varepsilon^0}$, $B_i$, $i = 1, ..., k$ are $p \times p$ matrices and $\mu^0$ is a $p \times 1$ vector; the elements in $B_i$ and $\mu^0$ depend on the elements in $\alpha$, $\Phi_i$s, and $\mu$. Furthermore, it can be proved that $B_k$ is restricted as

$$B_k = B_k^* \equiv \left[ \underset{p \times r}{B_{w1,k}} \vdots \underset{p \times (p-r)}{O} \right] \tag{17}$$

and that the VAR (16)-(17) is (asymptotically) stable, i.e. the roots of $\det(I_p - \sum_{i=1}^{k} B_i s^i) = 0$ are such that $\mid s \mid > 1$.

In practice, the definition of the vector $W_t$ and the specification of the stable VAR system in (16) requires either that the cointegration matrix $\beta$ is known, or that $\beta$ is replaced by a super-consistent estimate. In line with the Assumption (iii) of the paper, in Section 4 we shall replace $\beta$ with its super-consistent estimate based on the entire available sample.

As the VAR (16)-(17) is stationary by construction, we define the demeaned process $W_t^d = W_t - E(W_t)$, where $E(W_t) = B(1)^{-1}\mu^0$ and $B(L)W_t^d = \varepsilon_t^0$. Hereafter the analysis will be developed with respect to the $W_t^d$ process, in order to math the formulation (12) of the NKPC, which is also expressed in demeaned form. In this set-up, $\ell$-step ahead forecasts of $W_t^d$ can be computed as

$$\widehat{E}_{t-h} W_{t+\ell}^d = E(W_{t+\ell}^d | \mathcal{H}_{t-h}) = g_w' \ B^{\ell+h} \ \widetilde{W}_{t-h} \tag{18}$$

where $h$ is an integer, $\widetilde{W}_t = (W_t^{d\prime} \vdots W_{t-1}^{d\prime} \vdots ... \vdots W_{t-k+1}^{d\prime})'$ is the $pk \times 1$ state vector associated with the VAR (16),

$$B = \begin{bmatrix} B_1 & B_2 & \cdots & B_{k-1} & B_k^* \\ I_p & O & \cdots & O & O \\ \vdots & \ddots & & \vdots & \vdots \\ O & \cdots & & I_p & O \end{bmatrix} \tag{19}$$

is the $pk \times pk$ associated companion matrix, and $g_w$ is a $pk \times p$ selection matrix such that $g_w' \widetilde{W}_t = W_t^d$. The symbol '$\widehat{\ }$' above expectations in (18) is used to recall that in the present context the agents' forecasts do not necessarily coincide with the expectations taken under the REH.

From (18) it turns out that for $h = 1$ and $\ell = 1$ and $\ell = 0$, the forecasts of the variables in $W_t^d$ are given by the expressions

$$\widehat{E}_{t-1}\Delta\pi_{t+1}^d = g_\pi' \widehat{E}_{t-1} W_{t+1}^d = g_\pi' B^2 \, \widetilde{W}_{t-1} \tag{20}$$

$$\widehat{E}_{t-1}\Delta\pi_t^d = g_\pi' \widehat{E}_{t-1} W_t^d = g_\pi' B \, \widetilde{W}_{t-1} \tag{21}$$

$$\widehat{E}_{t-1}(\beta' Z_t)^d = g_{w_1}' \widehat{E}_{t-1} W_t^d = g_{w_1}' B \, \widetilde{W}_{t-1} \tag{22}$$

where $g_\pi$ is a selection vector such that $g_\pi' \widetilde{W}_t = \Delta\pi_t^d$, and $g_{w1}$ is a selection matrix such that $g_{w1}' \widetilde{W}_t = W_{1t}^d = (\beta' Z_t)^d$. Condition both sides of the NKPC (12) with respect to $\mathcal{H}_{t-1}$ and using the MDS property of $u_t^*$, yields the relation

$$g_\pi' \widehat{E}_{t-1} W_t^d = \psi \; g_\pi' \widehat{E}_{t-1} W_{t+1}^d + \omega \; g_{w_1}' \widehat{E}_{t-1} W_t^d \tag{23}$$

which by means of (20)-(22), and exploiting the fact that $\widetilde{W}_t \neq 0$ almost surely for each $t$, implies the following set of cross-equation restrictions

$$g_\pi' B (I_{pk} - \psi B) - \omega g_{w_1}' B = 0_{1 \times pk} \tag{24}$$

involving both the structural parameters $\psi$ and $\omega$, and the VAR coefficients $B_i$, $i = 1, 2, ..., k$.

It is worth noting that the restrictions in (24) generalize the approach originally proposed by Sargent (1979) and Campbell and Shiller (1987) for 'exact' rational expectations models (i.e. models not including a MDS disturbance). As stated in the Assumption (i), these restrictions coincide with the restrictions implied by REH only when the rational expectations solution of the NKPC (for a given process generating $x_t$) is nested within the VAR in (13). For this reason, it is in principle correct to disentangle between the REH and the VEH, although in the current literature the two notions are treated indiscriminately. In the rest of the paper we shall follow this tradition, expect where explicitly indicated.

Fanelli and Palomba (2007), Appendix, discuss in detail the nature of nonlinear restrictions of the form (24), and show the conditions under which the VAR in (16) is locally identifiable under these constraints; Fanelli (2008*a*), Appendix, focuses exactly on the restrictions (24), and show how these can be solved. Those authors prove that the cross-equation restrictions (24) can be opportunely transformed as explicit form constraints in which the VAR coefficients of one of the equations of the system (16) are expressed as unique nonlinear functions of $\psi$ and $\omega$, and the remaining VAR coefficients. The number of cross-equation restrictions is $n =$

$[p(pk-p+r)]-[(p-1)(pk-p+r)+2]=pk-p+r-2$, where the first term in square brackets is the number of free parameters of the unrestricted VAR (16)-(17), and the second term in square brackets is the number of free parameters (including $\psi$ and $\omega$) of the VAR (16)-(17) under the restrictions, see the Appendix in Fanelli (2008*a*) for details. The condition $pk \geq 3+(p-r)$ guarantees that $n \geq 1$, i.e. that the cross-equation restrictions are binding, hence testable. A likelihood-ratio test for the cross-equation restrictions implied by the NKPC can be computed by estimating the system (16)-(17) without (additional) restrictions, and subject to the cross-equation restrictions, and compared with $\chi^2_{n,1-\eta}$, where $\chi^2_{n,1-\eta}$ is the $1-\eta$ quantile of the $\chi^2$ distribution with $n$ degree of freedom, and $\eta$ is the level of the test.

### 3.2 Cross-equation restrictions under the ALH

When it is assumed that agents behave as econometricians and form their beliefs following adaptive learning rules, the analysis can be opportunely adapted. Since at time $t-1$ the VAR coefficients are not known and must be estimated from $\mathcal{H}_{t-1}$, in practice the agents' expectations are formed according to

$$\widehat{E}_{t-1}\pi^d_{t+1} = g'_\pi (B_{t-1})^2\, \widetilde{W}_{t-1} \tag{25}$$

$$\widehat{E}_{t-1}\pi^d_{t} = g'_\pi (B_{t-1})\, \widetilde{W}_{t-1} \tag{26}$$

$$\widehat{E}_{t-1}(\beta' Z_t)^d = g'_{w_1} (B_{t-1})\, \widetilde{W}_{t-1} \tag{27}$$

where with the notation $B_{t-1}$ we conventionally denote the counterpart of the companion matrix $B$ in (19), whose coefficients have to be replaced recursively, with the estimates based on $\mathcal{H}_{t-1}$.

Given initial coefficient estimates of the VAR (16)-(17) based on the sample from $t=1$ to $T_0$ ($\mathcal{H}_{T_0}$), imagine a situation in which at time $t=T_0+1$, $T_0+2$, ..., agents observe new data and update the estimates of VAR coefficients. In this case, the use of recursive least squares corresponds to the re-iterated application of Gaussian maximum likelihood estimation.[8] In this set-up also the cross-equation restrictions between the NKPC and the VAR coefficients must be updated and evaluated in correspondence of $t=T_0+1$, $T_0+2$, .... From (25)-(27), it turns out that the recursive counterpart of the restrictions in (24) take the form

$$g'_\pi B_{t-1}(I_{pk}-\psi B_{t-1}) - \omega g'_{w_1} B_{t-1} = 0_{1\times pk} \quad , \quad t=T_0+1, T_0+2, ... \tag{28}$$

where $T_0+1$ can be regarded as the first monitoring time, that means that the model will be recursively tested from $T_0+1$ onwards. Clearly, also the restrictions in (28) will admit, for each $t$, a unique explicit form representation, as it happens with the restrictions in (24).

[8] A widely used recursive updating rule in the adaptive learning literature is represented by a 'constant gain' version of recursive least squares (Sargent, 1999). This estimator discounts past observations at a geometric rate and is consequently more robust to structural changes (Branch and Evans, 2006) but is not consistent.

Consider now the objective of constructing a test for the null hypothesis that the sequence of cross-equation restrictions (28) implied by the NKPC hold, against the alternative that the VAR coefficients are unrestricted (in line with a backward-looking specification). For each given $t = T_0 + 1$, $T_0 + 2$ ,..., let

$$\log \widehat{L}_t^{\max} = -\frac{t}{2} \log(\det(\widehat{\Sigma}_{\varepsilon^0,t})) \tag{29}$$

be the log-likelihood (a part from a constant) of the VAR (16)-(17) evaluated at the maximum, where $\widehat{\Sigma}_{\varepsilon^0,t}$ is the estimated covariance matrix based on $\mathcal{H}_t$; let

$$\log \widetilde{L}_t^{\max} = -\frac{t}{2} \log(\det(\widetilde{\Sigma}_{\varepsilon^0,t})) \tag{30}$$

be the log-likelihood (a part from a constant) of the VAR (16)-(17) subject to the nonlinear cross-equation restrictions, where $\widetilde{\Sigma}_{\varepsilon^0,t}$ is the estimated covariance matrix of the restricted system based on $\mathcal{H}_t$. Given (29) and (30), likelihood-ratio statistics for the null hypothesis can be obtained through the sequence

$$LR_t = -2(\log \widetilde{L}_t^{\max} - \log \widehat{L}_t^{\max}) \quad , \quad t = T_0 + 1, T_0 + 2, ... \tag{31}$$

Although the ALH is logically based on a perpetual learning mechanism whose ultimate effect can be evaluated only in the limit, in practice the sequence (31) will be calculated over a finite number of periods, i.e. from $t = T_0 + 1$ until $t = T^{\max}$, where $T^{\max}$ is the length of the available sample at time the test is computed. The point is whether and how the observations from $T_0 + 1$ to $T^{\max}$ can be sensibly used to evaluate the NKPC under adaptive learning.

We refer to Fanelli (2008*b*) for a comprehensive discussion of sequences of recursively computed tests like (31). Using Monte Carlo experiments, Fanelli (2008*b*) shows that in order to control the size of such type of tests over the entire sequence $T_0 + 1$- $T^{\max}$, while preserving power against a backward-looking (unrestricted) VAR, a procedure which works successfully in finite samples is obtained by comparing each $LR_t$ in (31) with the critical value $cv_t$, where $cv_t$ is constructed as the linear combination $cv_t = a\chi^2_{n,1-\eta} + (1-a)IR_t^{n,\eta}$; here $\eta$ is the nominal size of the test, $\chi^2_{n,1-\eta}$ has been defined above, $IR_t^{n,\eta} = (c_{n,\eta})^2 + n\log(t/T_0)$ is the asymptotic critical value derived by Inoue and Rossi (2005) for sequences of recursive tests for predictive ability, $c_{n,\eta}$ can be taken from their Table 1 for given values of $n$ and $\eta$, and $a$ is a scalar, $0 < a < 1$, that must determined opportunely (see below).

The intuition behind the use of the critical value $cv_t$ in (31) is the following. If one compares the test statistics in the sequence (31) with the asymptotic critical value $\chi^2_{n,1-\eta}$, which do not vary over time, the resulting procedure is size-biased by construction; indeed, it can be proved that due to the law of iterated logarithms, the null hypothesis will be falsely rejected as the number of test repetitions increases over time (Robbins, 1970). The Monte Carlo experiments in Fanelli (2008$b$) show that for relatively small numbers of test repetitions, the empirical size of the test tend to depart considerably from the nominal one. On the other hand, if one compares the test statistics in the sequence (31) with the asymptotic critical value $IR_t^{n,\eta}$, which Inoue and Rossi (2005) have derived from Brownian motions for recursive tests for predictive ability, the resulting procedure turns out to be very conservative in finite samples, and displays low power against unrestricted VARs. For suitable choices of $a$, the critical value $cv_t$ defined above can be regarded as a 'scaled' version of the original $IR_t^{n,\eta}$, and exhibits a good compromise between the necessity of controlling the null hypothesis over the monitoring sequence $T_0+1$ - $T^{\max}$ (size), and the necessity of rejecting the model when the restrictions implied by the forward-looking model do not hold (power). In particular, the Monte Carlo experiments in Fanelli (2008$b$) show that in stationary VAR systems characterized by moderate values of $p$ and $k$, in which the number of structural parameters to estimate under the null hypothesis is small ($\psi$ and $\omega$ in the NKPC (12)), and considering samples of lengths typically available to researchers, the values $a = 0.75$ and $a = 0.50$ in $cv_t$ provide a satisfactory balance between size and power of the tests in the sequence.[9]

# 4 Results

We consider quarterly data relative to the euro area, using the last release of the Area-wide Model (AWM) data set described in Fagan et al. (2001). The variables used in the analysis cover the period 1980:4-2006:4. To measure inflation we use the GDP deflator, i.e. $\pi_t = 400 \times (p_t - p_{t-1})$, where $p_t$ is the log of the GDP deflator. As in Gali et al. (2001), firms' average marginal costs are proxied by the wage share (log of real unit labour costs), $x_t = ws_t$. A short term interest rate, $a_t = i_t$ ($q_a = 1$), has been included in the system, as interest rates can influence the marginal costs through the cost channel, see Chowdhury et al. (2005).

The plots of $\pi_t$ and $ws_t$ are reported in the upper panel of Figure 1, whereas in the lower panel we have plotted the short term nominal interest rate, $i_t$, and the real (ex-post) interest rate, $i_t - \pi_t$, respectively.[10] The simple graphical inspection seems to question the stationarity of the ex-post real interest rate in the euro area.

[9] An alternative route for recursively testing the NKPC under the ALH is proposed in Palomba and Fanelli (2007), who use simulation-based methods. Using the local Monte Carlo technique ('parametric bootstrap' or 'parametric Monte Carlo') formalized in Dufour and Jouini (2006), they compute the simulated p-value associated with each $LR_t$ in (31). As shown by those authors, this testing method has the advantage of working successfully in finite samples, but is computationally intensive in the adaptive learning framework.

[10] In the graph the inflation rate has been reported in the form $(p_t - p_{t-1})$, and the nominal interest rate has not been multiplied by 100.

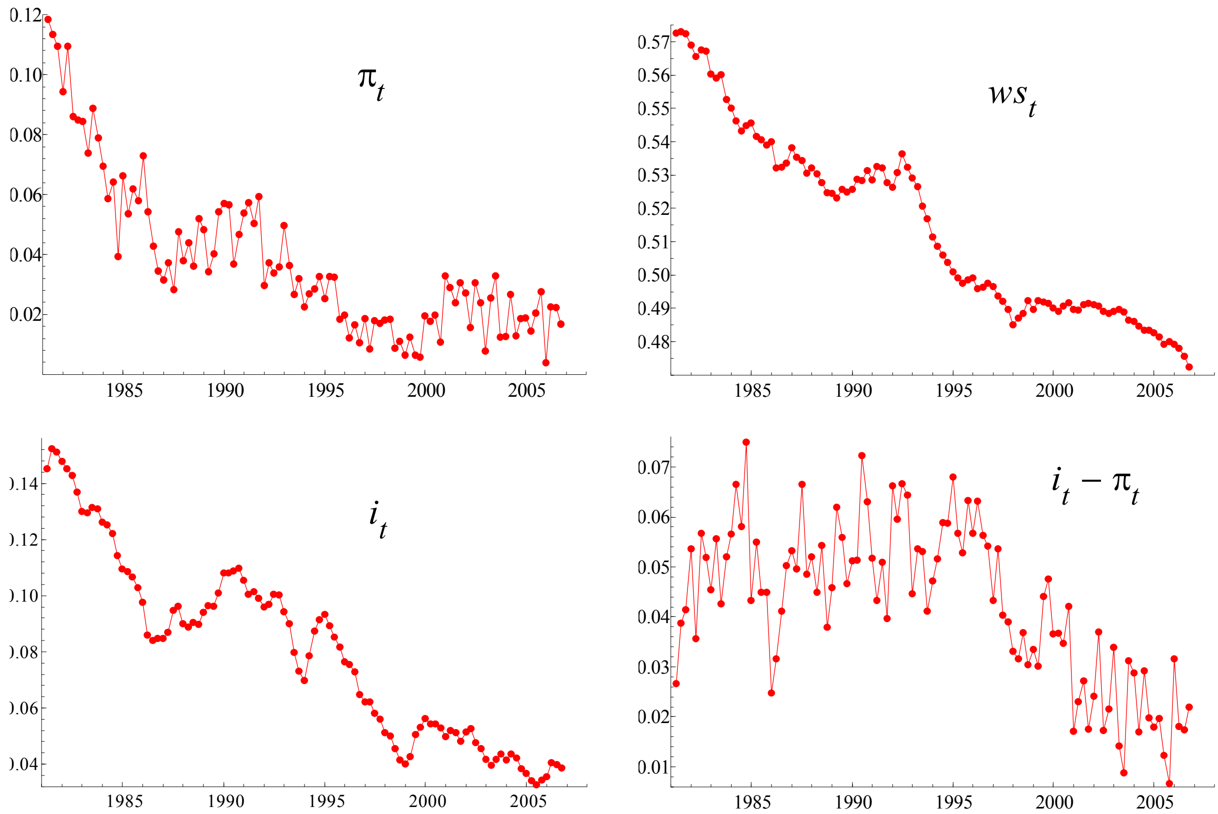


Figure 1. Euro area: inflation rate ($\pi_t$), wage share ($ws_t$), short-term nominal interest rate ($i_t$) and short-term real (ex-post) interest rate ($i_t - \pi_t$).

The analysis is initially based on a VAR for $Z_t = (\pi_t : ws_t : i_t)'$ with two lags ($k = 2$) and including an unrestricted constant. Conditional on intial values, the model is estimated over the period 1980:4-2006:4. Table 2 investigates the cointegration properties of the system. The likelihood ratio trace test (Johansen, 1996) suggests the presence of $p-r = 2$ common stochastic trends in system, though this evidence is not clear-cut. As it is known, the determination of the cointegration rank is a difficult choice in finite samples, and can be supported by other information in the model (Juselius, 2006). In the middle panel of Table 2 we have reported the inverse roots of the characteristic equation of the VAR, obtained in correspondence of the three possible values of $r$. It turns out that with $r = 0$ one should switch to a model in first differences alone to eliminate the unit roots in the system, complicating the analysis of the NKPC. On the other hand, with $r = 2$ (implying a single common stochastic trend) one root remains close to one, suggesting that imposing $r = 2$ and identifying one of the two cointegration vectors as a relation involving $i_t$ and $\pi_t$, would leave serious doubts about the stationarity of the identified relations. On the other hand, $r = 1$, suggested by the trace test, seems to remove all roots close to one from the system, and is consistent with a scenario where $i_t$ and $\pi_t$ do not cointegrate at all, see Figure 1.

VAR for $Z_t = (\pi_t{:}ws_t{:}i_t)'$ , $k = 2$ , 1981:2–2006:4

Cointegration rank test

| $H_0 : r \leq j$ | Trace | p-val |
|---|---|---|
| j=0 | 30.39 | 0.04 |
| j=1 | 11.51 | 0.18 |
| j=2 | 1.99 | 0.16 |

Roots of the system for $r = 0$: 0.986, 0.813, 0.667±0.241$i$, -0.391, 0.111

Roots of the system for $r = 1$: 0.623±0.105$i$, -0.396, 0.106

Roots of the system for $r = 2$: 0.815, 0.666±0.242$i$, -0.393, 0.118

Estimated cointegrating relation, $r = 1$

$$\widehat{\beta}' Z_t = W_{1t}^0 = \pi_t - \underset{(0.092)}{0.81}\ ws_t \quad , \text{LR: } \chi^2(1) = \underset{[0.07]}{3.29}$$

$$\widehat{\beta}' Z_t = W_{1t}^1 = ws_t \quad , \quad \text{LR: } \chi^2(2) = \underset{[0.000]}{16.68}$$

Table 2: Likelihood ratio trace test for cointegration rank, roots of the system for different values of cointegration rank, and estimated cointegrating relation for r=1. NOTES: Standard errors in parentheses, p-values in square brackets.

Thus, the choice $r = 1$ seems the best compromise between data properties and a priori expectations, and we have tested the two alternative formulations (5) and (8) of the NKPC in the lower panel of Table 2. The likelihood ratio tests for over-identifying restrictions on $\beta$ seem to favour form (8), thus rejecting the hypothesis of stationarity of the wage share. This results is in line with the findings in Fanelli (2008$a$) obtained on a previous release of the AWM data set, and on a different span of data. The estimated cointegration coefficient between $\pi_t$ and $ws_t$ is equal to $\widehat{\phi} = 0.81$.

Considering the VEqC representation of $Z_t = (\pi_t : ws_t : i_t)'$, and fixing $\beta$ at the super-consistent estimate of Table 2, such that $\widehat{\beta}' Z_t = \pi_t - \widehat{\phi} ws_t = \pi_t - 0.81ws_t$, we define the vector

$$W_t = \begin{pmatrix} \widehat{\beta}' Z_t \\ v'\Delta Z_t \end{pmatrix} = \begin{pmatrix} W_{1t} \\ W_{2t} \end{pmatrix} = \begin{pmatrix} \pi_t - 0.81ws_t \\ \Delta\pi_t \\ \Delta i_t \end{pmatrix} \tag{32}$$

where it can be recognized that $\det(v'\widehat{\beta}_\perp) \neq 0$, where $v' = (e_1' : e_3')$ and $e_i$ is a $p \times 1$ vector of zeros except in the $i$-th entry, which contains one. As shown in Section 3, $W_t$ in (32) admits a VAR representation of the form (16)-(17) with $k = 2$ lags and a constant. This VAR will be the statistical model upon which the NKPC (8) will be tested under the ALH.

Table 3 reports some vector residual diagnostic tests relative to the VAR for $W_t$ over the period 1981:2-2006:4, and remarks that the hypothesis of Gaussian uncorrelated disturbances can be taken as a reasonable approximation for this model. Moreover, the graphs in Figure 2, which report the one-step recursive residuals of the VAR with approximate 95% confidence bands computed over the period 1986:1-2006:4 (see below), support the findings in O'Reilly and Whelan (2005).

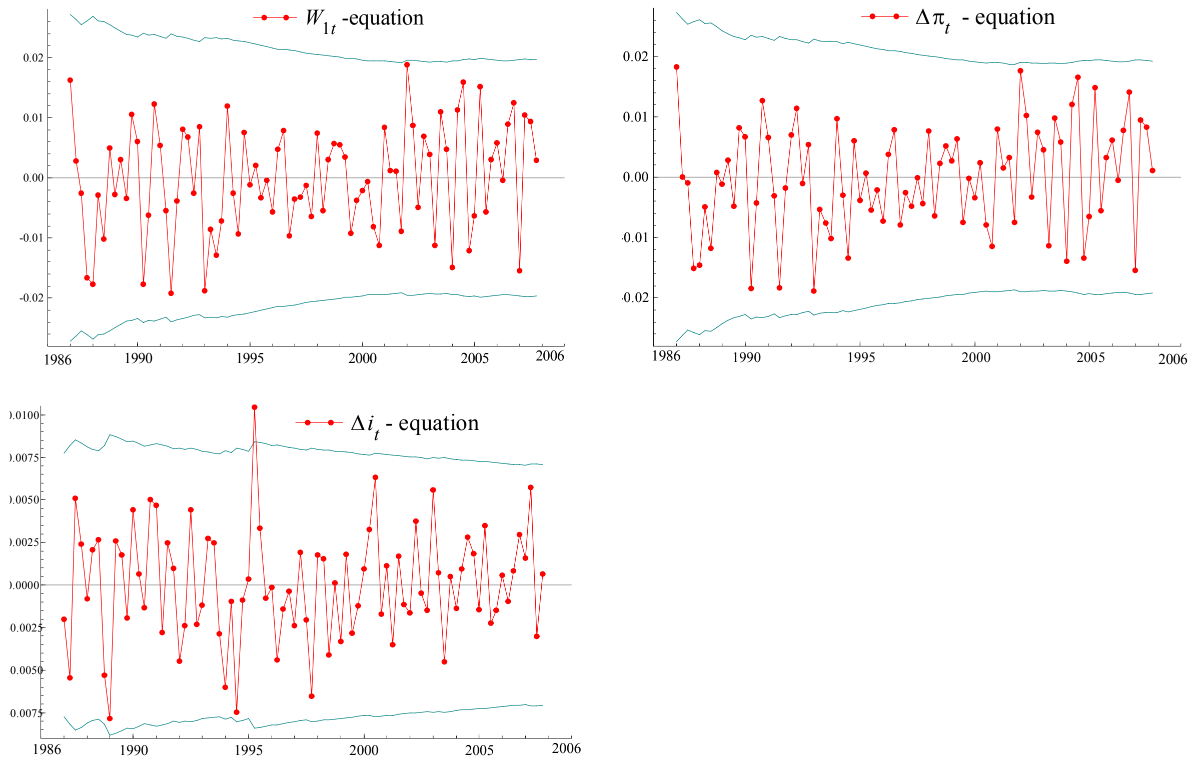


Figure 2. Euro area: one-step recursive residuals with approximate 95% confidence bands computed from the VAR for $W_t$ defined in (32) $(k = 2)$ over the period 1986:1-2006:4.

In order to compute likelihood-ratio tests for the cross-equation restrictions implied by the NKPC under the ALH, the VAR for $W_t$ must be estimated recursively both unrestrictedly and subject to the restrictions, for $t = T_0 + 1$, $T_0 + 2$,... We have chosen the sample 1986:1-2006:4 as the monitoring period; in terms of the notation of Section 3, $T_0 + 1$ =1986:1 and $T^{\max}$ =2006:4. The motivation for this choice is that after 1984/1985, the nature of the European Exchange Rate Mechanism characterizing the majority of European countries changed from a 'soft' to a 'hard' exchange rate parity arrangement (Batini, 2006); moreover, from 1986 onwards almost all European central banks adopted a more aggressive approach toward fighting inflation, with the objective of converging towards the European Monetary Union. Thus it can be argued that the econometric evaluation of the NKPC over the period 1986:1-2006:4 involves a relatively 'homogeneous' monetary policy regime. Accordingly, the VAR for $W_t$ has been fist estimated on the period 1981:1-1985:4 to form initial coefficients beliefs, and then has been estimated recursively over the period 1986:1-2006:4 to evaluate the restrictions implied by the NKPC, as detailed in Subsection 3.2.[11]

The maximization of the Gaussian likelihood of the VAR for $t = T_0 + 1, ..., T^{\max}$ is standard: as VAR coefficients are by construction subject to the zero constraints of the form (17), maximum likelihood estimation amounts to a feasible version of generalized least squares. The maximization of the constrained Gaussian likelihood of the VAR for $t = T_0 + 1, ..., T^{\max}$ has been performed by combining the BFGS method (Fletcher, 1987) with a grid search for the structural parameters $\psi$ and $\omega$.[12] Actually, we selected a grid of points for $\gamma_f$ and $\gamma_b$ (recall that here $\lambda = \widehat{\phi}(1 - (\gamma_f + \gamma_b)$ due to (11)) imposing the constraint $\gamma_f + \gamma_b < 1$, and then used the mapping (9)-(10) to recover the corresponding set of points for $\psi$ and $\omega$; $\gamma_f$ has been chosen in the range 0.6-0.98 and $\gamma_b$ in the range 0.0-0.30, using common incremental value 0.02.[13] In this case the number of cross-equation restrictions between the VAR coefficients and the NKPC is $n = pk - p + r - 2 = 2$.

Figure 3, uppuer panel, plots the sequence of likelihood ratio statistics computed over the monitoring period, along with the 5% ($\eta = 0.05$) critical values $\chi^2_{n,0.95}$, $IR_t^{n,\eta}$ and $cv_t = a\chi^2_{n,0.95}+ (1-a)R_t^{n,\eta}$, with $a$ equal to 0.75 and 0.50, respectively (see Subection 3.2). Figure 3, lower panel, reports the recursive maximum likelihood estimates of the structural parameters $\gamma_f$, $\gamma_b$ and $\lambda = \widehat{\phi}(1 - (\gamma_f + \gamma_b)$, obtained from the estimation of the constrained VAR.

[11] Carceles-Poveda and Giannitsarou (2007) show that the initialization issue may be relevant in finite samples; our testing results, however, are robust to different choices of $T_0 + 1$.

[12] Although the VAR includes a constant, the cross-equation restrictions have been derived considering only the coefficients in the matrices $B_i$, $i = 1, .., k$, see Subsection 3.1.

[13] In preliminary analysis, we used smaller values for $\gamma_f$ and larger values for $\gamma_b$ in the grid, without changes in the results. All computations have been performed through Ox 3.1.

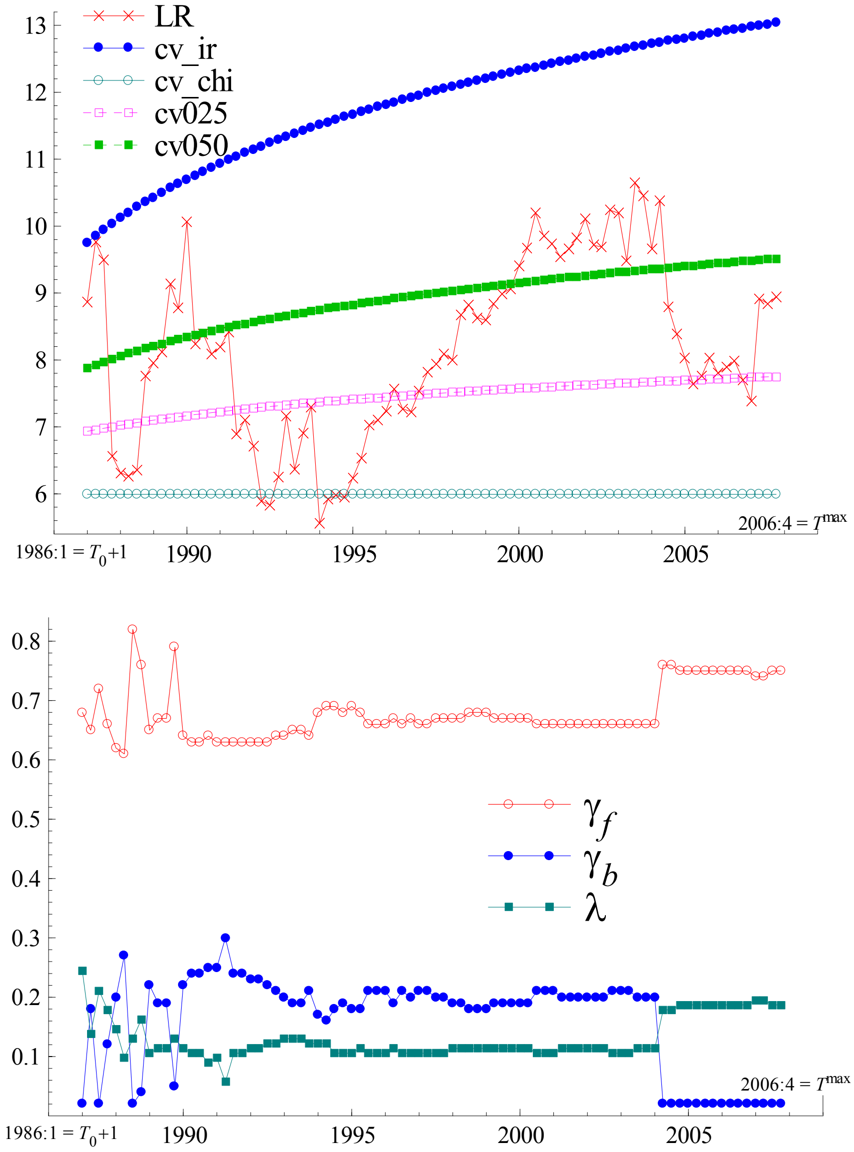


Figure 3. Euro area data. Upper panel: Sequence of recursively computed likelihood ratio (LR) statistics for the cross-equation restrictions implied by the cointegrated NKPC under the ALH, over the monitoring period 1986:1-2006:4, with corresponding 5% critical values (Section 3.2). Lower panel: recursively estimated structural parameters of the NKPC obtained through the grid-search. Results are here obtained through a VAR for the vector $W_t$ defined in (32), based on two lags ($k = 2$).

Figure 3 shows that if one considers a 'one-shot' test of the NKPC under the REH, namely compares the value of the likelihood-ratio statistic obtained at the date 2006:4 (using all available information from $t = 1$ until $t = T^{\max}$) with the quantile taken from the $\chi^2(2)$ distribution, the model is sharply rejected. This results is consistent with Bårdsen et al. (2004), O'Reilly and Whelan (2005), and Fanelli (2008), who provide some direct and indirect evidence against the NKPC. However, taking into account the sequential nature of the recursive test, and comparing each likelihood-ratio statistics in the sequence with the two types of critical values, $cv_t$, it turns out that the NKPC is supported about 30% of times using $a = 0.75$, and about 70% of times using $a = 0.50$.[14] As concerns the structural parameters reported in the lower panel of Figure 3, the magnitude of the estimated, $\gamma_f$, dominates the magnitude of the backward-looking parameter, $\gamma_b$, over the entire monitoring period; on the other hand, in this set-up the slope parameter $\lambda$ depends on $\gamma_f$ and $\gamma_b$ and the estimated cointegration parameter $\widehat{\phi}$, as implied by the relation (11). Referring to the 'primitive' parameters of the Calvo model, and using the mapping (2)-(4), the values of $\gamma_f$, $\gamma_b$ and $\lambda$ reported in Figure 3 imply that the discount factor, $\rho$, fluctuates in the range 0.76-0.79, that the average duration over which prices are kept fixed, $(1-\theta)^{-1}$, fluctuates in the range 3.3-3.5 (quarters), and that the fraction of backward-looking firms that change their prices following rule-of-thumb behavior, $\varpi$, fluctuates in the range 0.014-0.25, over the monitoring period. The dominance of the magnitude of $\gamma_f$ can be explained by observing that the agents' perceived law of motion, which enters the NKPC (12) through the term $E_t \Delta \pi^d_{t+1}$, already incorporates lagged values of the variables. This is a typical feature of adaptive learning algorithms.

In terms of robustness check, we repeated the analysis described above using a VAR for $Z_t = (\pi_t : ws_t : i_t)'$ with three lags ($k = 3$). Also in this case we have detected an identified single cointegration relation ($r = 1$) equivalent to the one reported in the last panel of Table 2. Accordingly, estimation and testing has been carried out using the VAR for $W_t$ defined in (32) but with three lags.[15] The number of cross-equation restrictions is $n = pk - p + r - 2 = 5$. Results have been reported in the two panels of Figure 4. It can be observed that in this case the NKPC is not rejected over the entire monitoring period using the two types of critical values $cv_t$; moreover, the sequence of estimated structural parameters appears more stable than before. Considering the parameters of the Calvo model through the mapping (2)-(4), we find that the values of $\gamma_f$, $\gamma_b$ and $\lambda$ reported in Figure 4 imply that the discount factor fluctuates in the range 0.80-0.82, the average duration over which prices are kept fixed fluctuates in the range 3.7-3.8 (quarters), and the fraction of backward-looking firms that change their prices following rule-of-thumb behavior fluctuates in the range 0.0148-0.015, over the monitoring period.

[14] We do not explicitly refer to the set of critical values $IR_t^{n,\eta}$ since, as argued in Subsection 3.2, in this case the test has low power against unrestricted VARs in finite samples.

[15] In terms of residuals diagnostic tests, the VAR with three lags fits the data almost as satisfactorily as the VAR with two lags investigated above, even if the latter performs better in terms of standard information criteria.

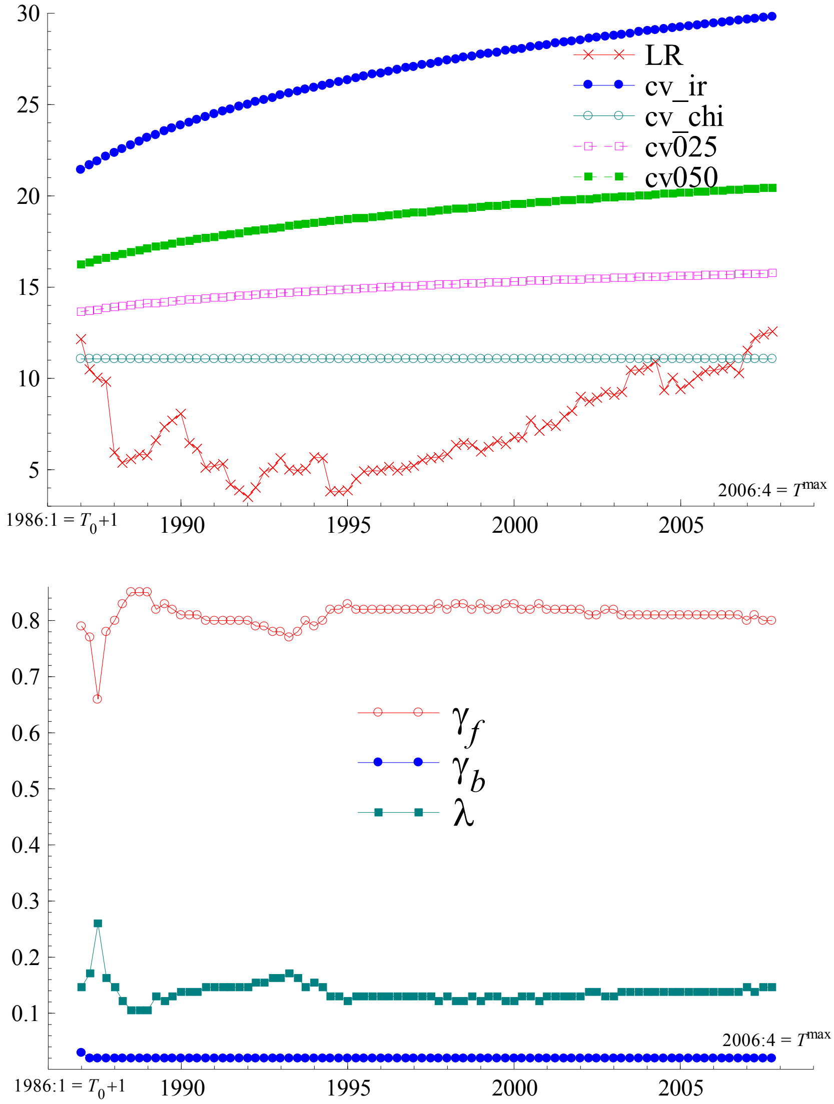


Figure 4. Euro area data. Upper panel: Sequence of recursively computed likelihood ratio (LR) statistics for the cross-equation restrictions implied by the cointegrated NKPC under the ALH, over the monitoring period 1986:1-2006:4, with corresponding 5% critical values (Section 3.2). Lower panel: recursively estimated structural parameters of the NKPC obtained through the grid-search. Results are here obtained through a VAR for the vector $W_t$ defined in (32), based on three lags ($k = 3$).

Overall, although the statistical evidence in support of the NKPC under the chosen adaptive learning algorithm is not striking, the results in this paper highlight that the NKPC might have some empirical support under the ALH. These findings can be related to other studies based on adaptive learning. Focusing on the U.S. economy, Milani (2005) considers a specification of the NKPC of the form (1), and a univariate autoregressive forecast model which is estimated recursively by using a constant gain version of recursive least squares (see footnote 8). He finds that when such a learning mechanism replaces the REH, structural sources of inflation persistence such as indexation are no longer essential to fit the data. Learning is thus interpreted as the major source of persistence in inflation. Differently from Milani (2005), our testing approach shows that adaptive learning behaviour does not suffice alone to explain the inertia in the data captured by the lagged inflation term in (1), i.e. $\gamma_b \neq 0$. This evidence is consistent with both Fanelli and Palomba (2007) and Fanelli (2008*b*), who also investigate the NKPC attempting to test the VAR implications of the ALH.

Finally, the results of the paper can be also reconciled with Weber (2007), who using survey data of household and expert inflation expectations, finds that for major European countries inflation expectations and are not rational, and rather result from a learning process.

| | VAR for $W_t = (W_{1t}{:}\Delta\pi_t{:}\Delta i_t)'$ , $k = 2$ , 1981:2-2006:4 |
|---|---|
| Autocor. | F(45,241)=1.39 [0.06] |
| Normality | $\chi^2(6)$=0.60 [0.36] |
| Roots | 0.623$\pm$0.105$i$ , -0.396, 0.106 |

Table 3: Vector residual diagnostic tests for the VAR defined in (32). NOTES: Autocor= LM vector test for residual autocorrelations up to 5 lags; Normality = LM vector test for residual normality; p-values in square brackets; Roots = inverse roots of the characteristic equation of the VAR.

## 5 Concluding remarks

In this paper we have provided a method for evaluating the data adequacy of the NKPC under a formulation of the ALH which extends the econometrics analysis based on the VEH to a recursive framework. The idea is that agents use VAR models as their perceived law of motion, and update coefficients recursively through maximum likelihood estimation, as new data enter the information set. The ALH gives rise to a sequence of cross-equation restrictions between the VAR coefficients and the NKPC, rather than a single set of constraints. The cross-equation restrictions can be recursively tested by computing sequences of likelihood-ratio statistics over the monitoring period, but standard (time invariant) critical values can not be applied. Insights on how reliable inference can be applied in these circumstances have been provided. The method has been developed by focusing on the case of cointegrated variables, and it has been assumed that the agents' adaptive learning rule involves only the short run adjustment coefficients of the system, and not the cointegration parameters. This assumption can be relaxed in future research, and potentially alternative recursive estimation methods can be applied.

The empirical analysis based on quarterly data for the euro area has shown that the inflation rate and the wage share can be approximated as nonstationary cointegrated processes over the period 1981-2006. Imposing the cointegration restrictions and using the VAR as the statistical platform upon which the restrictions implied by the NKPC have been tested, it turns out that there is little room for the NKPC under the REH. On the other hand, assuming that the agents deviate from the REH, use a VAR for the data to form their forecasts, and learn gradually the parameters of the model as they face new data, we have found that there are pieces of evidence in support of the model over the period 1986-2006. Moreover, we have argued that typical claims about the ability of adaptive learning rules to replace inertial inflation mechanisms may be misleading, if not properly tested.

Overall, the evidence presented in this paper suggests that if one firmly believes in the NKPC, then the data tend to favour the restrictions implied by the ALH over the restrictions implied by the REH. In line with the objectives of the paper, however, we have not investigated whether apart from the agents' expectations generating mechanism, alternative sources of price growth other than marginal costs contain considerable information on inflation dynamics in the euro area.

## References

Bårdsen, G., Jansen, E. S. and Nymoen, R. (2004), Econometric evaluation of the New Keynesian Phillips curve, *Oxford Bulletin of Economics and Statistics* 66 (Supplement), 671-685.

Batini, N. (2006), Euro area inflation persistence, *Empirical Economic* 31, 977 - 1002.

Boug, P., Cappelen, and Swensen, A. R. (2007), The New Keynesian Phillips Curve revisited, *Statistics Norway, Discussion Paper* No. 500.

Branch, W.A. (2004), The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations, *Economic Journal* 114, 592-621.

Branch, W.A., Evans, G.W. (2006), A simple recursive forecasting model, *Economic Letters* 91, 158-166.

Brayton, F., Eileen, M., Reifschneider, Tinsley, P., Williams, J. (1997), The role of expectations in the FRB/US macroeconomic model, *Federal Reserve Bulletin* 83 (April), 227-245.

Bullard, J., Mitra, K. (2007), learning about monetary policy rules, *Journal of Monetary Economics* 49, 1105-1129.

Campbell, J. Y., Shiller, R. J. (1987), Cointegration and tests of present value models, *Journal of Political Economy* 95, 1062-1088.

Carceles-Poveda, E., Giannitsarou, C. (2007), Adaptive learning in practice, *Journal of Economic Dynamics and Control* 31, 2659-2697.

Chowdhurya, I., Hoffmann, M., Schabertb, A. (2005), Inflation dynamics and the cost channel of monetary transmission, *European Economic Review* 50, 995-1016.

Dees, S., Pesaran, M.H., Snith, V.L., Smith, R.P. (2008), Identification of New Keynesian Phillips curves from a global perspective, *CESifo* Working Paper Series No. 2219.

Dufour, J.-M, Jouini, T. (2006), Finite-sample simulation-based inference in VAR models with application to Granger causality testing, *Journal of Econometrics* 135, 229-254.

Evans, G. W. and Honkapohja, S. (1999), *Learning dynamics*, in Handbook of Macroeconomics 1A, Chap. 7

Evans, G. W., Honkapohja, S. (2001), *Learning and expectations in macroeconomics*, Princeton University Press.

Evans, G. W., Honkapohja, S. (2003*a*), Adaptive learning and monetary policy design, *Journal of Money, Credit and Banking* 35, 1045-1072.

Evans, G. W., Honkapohja, S. (2003*b*), Expectations and the stability problem for optimal monetary policies, *Review of Economic Studies* 70, 807-824.

Evans, G. W., McGough, B. (2005), Monetary policy, ideterminacy and learning, *Journal of Economic Dynamics and Control* 29, 1809-1840.

Fagan, G., Henry, G. and Mestre, R. (2001), An area-wide model (awm) for the Euro area, *European Central Bank, Working Paper* No. 42.

Fanelli, L. (2008*a*), Testing the New Keynesian Phillips Curve through Vector Autoregressive models: Results from the Euro area, *Oxford Bulletin of Economics and Statistics* 70, 53-66.

Fanelli, L. (2008*b*), Likelihood-based recursive tests of the adaptive learning hypothesis, mimeo, Working Paper, available at: http://www.rimini.unibo.it/fanelli/recursive_test.pdf

Fanelli, L., Palomba, G. (2007), Simulation-based tests of forward-looking models under VAR learning dynamics, *Universita' Politecnica delle Marche, Quaderno di Ricerca* No. 298, http://dea.univpm.it/quaderni/pdf/298.pdf.

Fletcher, R., (1987), *Practical methods of optimization*, Wiley-Interscience, New York.

Fuhrer, J. (1997), The (un)importance of forward-looking behavior in price specifications, *Journal of Money Credit and Banking* 29, 338-350.

Fuhrer, J., Moore, G. (1995), Inflation persistence, *Quarterly Journal of Economics* 110, 127-159.

Galí, J., Gertler, M. (1999), Inflation dynamics: a structural econometric analysis, *Journal of Monetary Economics* 44, 195-222.

Galí, J., Gertler M. and Lopez-Salido, J.D. (2001), European inflation dynamics, *European Economic Review* 45, 1237-1270.

Goldberg, M., Frydman, R. (1996), Imperfect knowledge and behaviour in the foreign exchange market, *Economic Journal* 106, 869-893.

Goldberg, M., Frydman, R. (2007), *Imperfect knowledge economics*, Princeton University Press, Princeton.

Hendry, D.F. (1988), The encompassing implications of feedback versus feedforward mechanisms in econometrics, *Oxford Economic Papers* 40, 132-149.

In-Koo, C., Kasa, K. (2005), Learning and model validation, mimeo.

Inoue, A., Rossi, B. (2005), Recursive predictability tests for real-time data, *Journal of Business and Economic Statistics* 23, 336-345.

Johansen, S. (1996), *Likelihood-based inference in cointegrated Vector Auto-Regressive models*, Oxford University Press, Second revised version, Oxford.

Johansen, S. (2006), Confronting the economic model with the data, in Colander, D. (ed): *Post Walrasian Macroeconomics*, Cambridge University Press, Cambridge, pp. 287-300.

Johansen, S., Swensen, A. R. (1999), Testing exact rational expectations in cointegrated vector autoregressive models, *Journal of Econometrics* 93, 73-91.

Juselius, K. (2006), *The cointegrated VAR model*, Oxford University Press, Oxford.

Juselius, M. J. (2006), Testing the new Keynesian model on US and European data, *61st European Meeting of the Econometric Society.*

Juselius, K., Franchi, M. (2007), Taking a DSGE model to the data meaningfully, *Economics Discussion Papers*, No 2007-6. http://www.economics-ejournal.org/economics/discussionpapers/2007-6

Kozicki, S., Tinsley, P. A. (1999), Vector rational error correction, *Journal of Economic Dynamics and Control* 23, 1299-1327.

Kurmann, A. (2007), Maximum likelihood estimation of dynamic stochastic theories with an application to New Keynesian pricing, *Journal of Economic Dynamics and Control* 31, 767-796.

Lindé, J. (2005), Estimating New Keynesian Phillips curves: A full information maximum likelihood approach, *Journal of Monetary Economics* 52, 1135-1149.

Lubik, T., Schorfheide, F. (2004), Testing for indeterminacy: an application to U.S. monetary policy, *American Economic Review* 2004, 94, 190-217.

McCallum, B. T. (1983), On non-uniqueness in rational expectations models: an attempt at perspective, *Journal of Monetary Economics* 11, 139-168.

McCallum, B. T. (2003), The unique minimum state variable RE solutuion is E-stable in all well formulated models, *NBER, Working Paper* No. 9960.

Mavroeidis, S. (2005), Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve, *Journal of Money Credit and Banking* 37, 421-448.

Milani, F. (2005), Adaptive learning and inflation persistence, *Working Paper*, University of California, Irvine.

Muth, J.F. (1961), Rational expectations and the theory of price movements, *Econometrica* 29, 315-335.

Nason, J. M., Smith, G. W. (2005), Identifying the New Keynesian Phillips curve, *Queen's Economics Department, Working Paper* No. 1026.

Orphanides, A., Williams, J.C. (2005), Monetary policy and learning, *Review of Economic Dynamics* 8, 498-527.

O'Reilly, G., Whelan, K. (2005), Has euro-area inflation persistence changed over time?, *Review of Economics and Statistics* 87, 709-720.

Paruolo, P. (2003), Common dynamics is I(1) VAR systems, *Università dell'Insubria, Quaderno di Ricerca* No. 2003/35, http://eco.uninsubria.it/dipeco/Quaderni/files/QF2003_35.pdf

Pesaran, H. M. (1987), *The limits to rational expectations*, Basil Blackwell, Oxford.

Robbins, , H. (1970), Statistical methods related to the law of the iterated logarithm, *Annals of Mathematical Statistics* 41, 1397-1409.

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

Ruud, J., Whelan, K. (2005*a*), Does labor's share drive inflation ?, *Journal of Money Credit and Banking* 37, 297-312.

Ruud, J., Whelan, K. (2005*b*), New tests of the New Keynesian Phillips curve, *Journal of Monetary Economics* 52, 1167-1181.

Ruud, J., Whelan, K. (2006), Can rational expectations sticky-price models explain inflation dynamics?, *American Economic Review* 96, 303-320.

Sargent, T.J. (1979). A note on the maximum likelihood estimation of the rational expectations model of the term structure, *Journal of Monetary Economics* 5, 133-143.

Sargent, T.J. (1999), *The conquest of American inflation*, Princeton University Press, Princeton.

Weber, A. (2007), Heterogeneous expectations, learning and European inflation dynamics, *Deutsche Bundesbank, Discussion Paper* No. 16/2007.