Università Politecnica delle Marche

DIPARTIMENTO DI ECONOMIA

Simulation-Based Tests of Forward-Looking Models Under VAR Learning Dynamics

Luca Fanelli Giulio Palomba

QUADERNI DI RICERCA n. 298

Settembre 2007

Abstract

In this paper we propose simulation-based techniques to investigate the finite sample performance of likelihood ratio (LR) tests for the nonlinear restrictions that arise when a class of forward-looking (FL) models, typically used in monetary policy analysis, is evaluated with Vector Autoregressive (VAR) models. We consider both 'one-shot' tests and sequences of tests under a particular form of adaptive learning dynamics, where 'boundedly rational' agents use VARs recursively to update their beliefs. The analysis is based on the comparison of the likelihood of the unrestricted and restricted VAR, and the p-values associated with the LR statistics are computed by Monte Carlo simulation. We also address the case where the variables of the FL model are approximated as non-stationary cointegrated processes. Application to the New Keynesian Phillips Curve in the euro area shows that the FL model of inflation dynamics is not rejected once the suggested simulation-based tests are applied. The result is robust to specification of the VAR as a stationary (albeit highly persistent) or cointegrated system. However, in the second case the imposition of cointegration restrictions changes the estimated degree of price stickiness.

JEL Class.: C32, C12, C52, D83, E10

Keywords: Adaptive learning, Cross-equation restrictions, Forward-looking model, Monte Carlo test, New Keynesian Phillips Curve, Simulation techniques, VAR.

Indirizzo: Luca Fanelli (corresponding author), Dipartimento di Scienze Statistiche, Università di Bologna, via Belle Arti n. 41, I-40126 Bologna, Italy, tel. +39 0541 434303, fax: +39 051 232153, *luca.fanelli@unibo.it*.
Giulio Palomba, Dipartimento di Economia, Università Politecnica delle Marche, Piazzale Martelli n. 8, I-60121 Ancona, Italy. *g.palomba@univpm.it*

Simulation-Based Tests of Forward-Looking Models Under VAR Learning Dynamics*

Luca Fanelli *Giulio Palomba*

1 Introduction

The class of forward-looking (FL) models typically employed in macroeconomics and monetary policy analysis, imposes parametric restrictions on vector autoregressive (VAR) systems for the variables, see, *inter alia*, Sargent (1979), Campbell and Shiller (1987), Baillie (1989) and Johansen and Swensen (1999). In many cases, these restrictions are highly nonlinear, as shown in Pesaran (1987), Binder and Pesaran (1995), Bekaert and Hodrick (2001) and Fanelli (2007b).

Tests in VAR models are usually based on linear restrictions and large-sample approximations; however, it is well recognized that the use of asymptotic distributions can lead to misleading inference, given the usual sample lengths available to macroeconometricians. Furthermore, the presence of non-stationary such as integrated processes can worsen reliability problems, as documented in Toda and Yamamoto (1995), Johansen (1996) and Yamada and Toda (1998). Although a great body of literature is currently devoted to envisaging the finite sample performance of tests for linear restrictions (Dufour and Jouini, 2006), less attention has been devoted to nonlinear restrictions, and in particular to the constraints that arise under rational (model consistent) expectations (RE). It is well known that RE impose tight restrictions on the models describing the data such as the VAR, and that these restrictions are usually rejected, suggesting that FL models are too simple to capture the complex probabilistic nature of the data.[1]

The contributions of this paper are twofold. First, compared to the existing literature where the finite sample evidence associated with linear (exclusion) restrictions, such as Granger non-causality constraints, is explored in VAR systems, we extend the analysis to the nonlinear restrictions that arise when FL models are tested under 'VAR expectations'. This type of expectations does not necessarily coincide with RE (Brayton *et al.*, 1997), however, its use in the

* *We wish to thank Hashem Pesaran and Riccardo "Jack" Lucchetti for helpful comments and suggestions on earlier drafts of the paper. We are responsible for all remaining errors.*

[1] For this reason, many authors argue that the empirical evaluation of FL models (as well as dynamic stochastic general equilibrium models) must be based on alternative criteria, including Bayesian techniques, see Canova (2007).

econometric practice is widespread and goes back to Sargent (1979), Hansen and Sargent (1980) and Campbell and Shiller (1987). We investigate whether the rejection of FL models can also be ascribed to the poor finite sample performances of the employed tests, such as e.g. likelihood ratio (LR) tests. Indeed, appealing to asymptotic critical values with relatively small samples might induce over-rejection of the FL model; we address the issue through simulation-based inference. Observe that we rely on LR tests rather than Wald tests, because the latter are known to possess bad small sample properties in models involving forward-looking behaviour, see Bekaert and Hodrick (2001) for a comprehensive discussion.[2]

Second, we develop a method to analyze the finite sample performances of sequences of LR tests for FL models that arise under a particular form of adaptive learning dynamics, where VARs serve as agents' perceived law of motion (PLM). The idea is that at each time t, agents use a VAR to form their beliefs using the available information;[3] as the information set increases, agents update their estimates through recursive methods, hence face a sequence of cross-equation restrictions in correspondence with each t.

When dealing with sequences of LR statistics, conventional critical values which do not take the recursive nature of the test into account are not suited; indeed, by the law of iterated logarithms, the probability of rejecting the null is asymptotically one (see Inoue and Rossi, 2005). Moreover, in this framework, the empirical assessment of the FL model involves relatively small sample sizes, especially at the beginning of the agents' learning process, challenging the reliability of asymptotic theory. Also in this case finite sample simulation-based techniques can deliver reliable inference.

The idea of the paper is to use the consistent (point) estimates of VAR parameters to implement local Monte Carlo (LMC) LR-type tests described in Dufour (2006) and Dufour and Jouini (2006). In practice, LMC amounts to a 'parametric bootstrap' (or 'parametric Monte Carlo') procedure, through which the p-value associated with the LR statistic can be calculated by simulation.

We consider two possible approaches to the empirical evaluation of FL models. In the former, we address the situation in which the researcher's objective is simply testing the VAR coefficient restrictions, regardless of the possible presence of unit roots. In the latter, it is explicitly recognized that in many circumstances the variables of the FL model might be approximated as non-stationary cointegrated processes, hence a pre-test for unit roots and cointegration is implicitly

[2]Clearly, we do not use a GMM approach since the objective of the paper is testing the cross-equation restrictions implied by FL models in a 'full-information' framework.

[3]The traditional approach to modelling 'boundedly rational' expectations assumes agents form expectations by using adaptive updating rules, see *e.g.* Branch and Evans (2006). We refer to *e.g.* Pesaran (1987), Ch. 3, Sargent (1999), and Evans and Honkapohja (1999, 2001) (see also references therein) for a comprehensive discussion of the concepts of 'bounded rationality' and learning behaviour.

assumed.

This paper provides an empirical illustration based on the New Keynesian Phillips Curve (NKPC) for the euro area, using a VAR for the inflation rate, the wage share and the interest rate to proxy agents' expectations. Our results show that the NKPC cross-equation constraints are not rejected when the LMC LR-type tests are computed, while they are rejected using the standard chi-squared distribution; this finding supports the idea that, in finite samples, tests based on asymptotic critical values may falsely lead one to reject the FL model. Furthermore, the estimated forward-looking parameter dominates in magnitude the estimated backward-looking parameter, indicating that learning dynamics might represent a potential source of euro area inflation persistence (Milani, 2005 for the US economy, and Fanelli, 2007a for the euro area).

The empirical analysis also highlights that detecting the presence/absence of unit roots and cointegration in the VAR before estimating and testing the FL model is an important aspect of the empirical modelling strategy. Indeed, if the inflation rate and the wage share are modelled as I(1) cointegrated processes, the estimated magnitude of the parameter that governs the pass-through of marginal costs into inflation reflects an higher degree of price stickiness, compared to the case where time-series are treated as stationary processes.

The paper is organized as follows. Section 2 introduces the FL model and derives the cross-equation restrictions with the VAR. Section 3 extends the analysis to the case where the variables of the system are cointegrated. Section 4 describes how the simulation-based LMC LR-type tests work with FL models, and Section 5 deals with a simple MC experiment. Section 6 consists of an empirical illustration relative to the NKPC for the euro area. Proofs and technical details are summarized in the Appendix.

2 Forward-looking model and VAR restrictions

We focus on the following class of FL models

$$\widetilde{y}_t = \gamma E_t \widetilde{y}_{t+1} + \delta y_{t-1} + \kappa \, \widetilde{w}_t + \rho_t \tag{1}$$

$$\widetilde{w}_t = m(\widetilde{y}_{t-1}, \widetilde{w}_{t-1}, ...) + e_t \ , \tag{2}$$

where $\widetilde{y}_t$ is a scalar endogenous variable, $E_t \widetilde{y}_{t+1} = E(\widetilde{y}_{t+1} \mid \mathcal{F}_t)$ is the expected value of $\widetilde{y}_{t+1}$ conditional on the sigma-field $\mathcal{F}_t$, $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$, $\widetilde{w}_t$ is an explanatory (driving) variable, $m(\cdot)$ is a linear function whose arguments are a finite number of lags of $\widetilde{y}_t$ and $\widetilde{w}_t$, respectively, associated with a given set of parameters, $\xi_t = (\rho_t, e_t)'$ a 2×1 martingale difference sequence (MDS) with respect to $\mathcal{F}_{t+1}$, and $\zeta = (\gamma,\delta,\kappa)'$, $0 < \gamma < 1$, $0 < \delta < 1$, $\kappa > 0$, is the vector of structural parameters.

The specification (1)-(2) covers many of the FL models typically used in monetary policy analysis, see Section 6. In this context, the symbol '$\widetilde{\ }$' over variables denotes that time-series do not embody any deterministic component. We have deliberately left the equation for $\widetilde{w}_t$ in (2) not fully unspecified; knowledge of the structure of the process generating $\widetilde{w}_t$ is crucial to compute the RE solution of the model and to envisage whether the model has a determinate (unique, non-explosive) solution or not, with consequences on the identifiability of structural parameters, see Mavroeidis (2005) for a comprehensive discussion. Such an approach, which is widespread in the literature, is based on the maintained assumption that the data generating process (DGP) belongs to the RE solution of the FL model.

Our approach, based on 'VAR expectations', takes a different perspective. It posits that the agents in the economy form their expectations from a VAR serving as the statistical forecast model for the data they face. The maintained assumption is that if such a reduced form is correctly specified, in the sense of capturing the dynamics and persistence of variables sufficiently well, the DGP must belong to the VAR, irrespective of whether the theory holds or not; consequently, if for a given $m(\cdot)$ the FL model (1)-(2) really holds, its (unique) RE solution must be nested within the statistical model. The method of undertermined coefficients allows to retrieve the implied cross-equation restrictions.

Let $X_t = (y_t, w_t, u_t')'$ be the $p \times 1$ vector of observable variables, where u_t is an $q \times 1$ ($p = 2 + q$) sub-vector of 'additional' variables that might enter agents' information set. The vector X_t collects the variables whose dynamics is potentially affected by deterministic components (constant, trend, dummies, etc.); $\widetilde{X}_t = (\widetilde{y}_t, \widetilde{w}_t, \widetilde{u}_t')'$ is the corresponding vector of variables net of determinist components, hereafter called the 'detrended' process.

Given the sigma field $\mathcal{H}_t = \sigma(X_1, ..., X_t) \subseteq \mathcal{F}_t$, it is assumed that X_t is generated by the finite-order VAR processes of the form

$$X_t = \widetilde{X}_t + d_t \quad , \quad A(L)\widetilde{X}_t = \varepsilon_t \quad , \quad A(L)d_t = \mu + \Theta D_t \tag{3}$$

where

$$A(L) = I_p - A_1 L - \cdots - A_k L^k,$$

L is the lag operator, A_i, $i = 1, 2, ..., k$ are $p \times p$ matrices of parameters, $k \geq 2$, d is a $p \times 1$ constant, D_t is a $f \times 1$ vector containing other deterministic terms (trend, dummies, etc.) with associated $p \times f$ matrix of parameters, Θ, and ε_t is a MDS with respect to $\mathcal{H}_t$, with $p \times p$ covariance matrix Σ_ε, and Gaussian distribution. The quantities X_0, X_{-1}, ..., X_{-1+k} are fixed. It is further assumed that the roots, z, of $\det[A(z)] = 0$ are such that $\mid z \mid \geq 1$, hence explosive solutions are ruled out.

Henceforth the VAR system in equation (3) will read as agents' expectations generating system, or PLM. This means that the agents in the economy use (3) and the information set $\mathcal{H}_t$ to form their beliefs over time. Since $\widetilde{X}_t = X_t - d_t = (\widetilde{y}_t, \widetilde{w}_t, \widetilde{u}_t')'$ is expressed as $A(L)\widetilde{X}_t = \varepsilon_t$, the first-order companion form representation of the detrended process is given by

$$X_t^* = AX_{t-1}^* + \varepsilon_t^*$$

where $X_t^* = (\widetilde{X}_t', \widetilde{X}_{t-1}', \cdots, \widetilde{X}_{t-k+1})'$ and $\varepsilon_t^* = (\varepsilon_t', 0_{1\times p(k-1)})'$ are $pk \times 1$, and A is the $n \times n$ companion matrix

$$A = \begin{bmatrix} A_1 & \cdots & A_k \\ I_p & & 0_{p\times p} \\ & \ddots & \\ 0_{p\times p} & \cdots & I_p \end{bmatrix}.$$

At time $t-1$, agents' j-step ahead forecast of $\widetilde{y}_t$ and $\widetilde{w}_t$, net of deterministic components, are given by

$$E(\widetilde{y}_{t+j} \mid \mathcal{H}_{t-1}) = s_y' E(X_{t+j}^* \mid \mathcal{H}_{t-1}) = s_y' A^{j+1} X_{t-1}^* \tag{4}$$
$$E(\widetilde{w}_{t+j} \mid \mathcal{H}_{t-1}) = s_w' E(X_{t+j}^* \mid \mathcal{H}_{t-1}) = s_w' A^{j+1} X_{t-1}^* \tag{5}$$

where s_y and s_w are respectively $pk \times 1$ selection vectors such that $s_y' X_t^* = y_t$ and $s_w' X_t^* = w_t$.

Using the law of iterated expectations ($\mathcal{H}_{t-1} \subseteq \mathcal{H}_t \subseteq \mathcal{F}_t$) and the MDS property of ρ_t, the FL equation (1) can be re-written as

$$E(\widetilde{y}_t \mid \mathcal{H}_{t-1}) = \gamma E(\widetilde{y}_{t+1} \mid \mathcal{H}_{t-1}) + \delta \widetilde{y}_{t-1} + \kappa\, E(\widetilde{w}_t \mid \mathcal{H}_{t-1})$$

so that using the forecasts (4)-(5) one has

$$s_y' A X_{t-1}^* = \gamma s_y' A^2 X_{t-1}^* + \delta s_y' X_{t-1}^* + \kappa s_w' A X_{t-1}^*.$$

As $X_t^* \neq 0$ a.s. for each t, rearranging terms above gives rise to the following set of nonlinear cross-equation restrictions

$$s_y' A(I_{pk} - \gamma A) - \delta s_y' - \kappa s_w' A = 0_{1\times pk}. \tag{6}$$

The two propositions that follow discuss the conditions which ensure that the VAR in equation (3) is locally identifiable under the restrictions (6), and the explicit form representation of these restrictions, respectively.

Proposition 1 *Given the VAR system (3) and the cross-equation restrictions (6), let $a_y' = s_y' A = (a_{y,1}, a_{y,2}, ..., a_{y,pk})$, $a_w' = s_w' A = (a_{w,1}, a_{w,2}, ..., a_{w,pk})$ and $a_u = vec(A_u)'$, $A_u = s_u' A$ be the $1 \times pk$, $1 \times pk$ and $1 \times qpk$ vectors containing the parameters associated with the y_t, w_t and u_t marginal equations of the VAR, respectively, with s_u' a $q \times n$ selection matrix such that $s_u' X_t^* = u_t$. Define the $[pk(q+1)+3] \times 1$ vector $v = (a_y', a_u', \zeta)'$. If $pk \geq 4$ and $a_{y,2} \neq -(\kappa/\gamma)$ in a* *neighbourhood of 'true' parameter values, then the following conditions hold: (i) the restrictions (6) can be uniquely expressed in the form*

$$a_w = g(v) \tag{7}$$

where $g : \mathbb{T} \to \mathbb{R}^{pk}$ is a differentiable function with $\mathbb{T}$ open set of $\mathbb{R}^{pk(q+1)+3}$; (ii) $c = pk - \dim(\zeta)$ is the number of over-identifying restrictions; (iii) the information matrix of the VAR under the restrictions (7) is non-singular.

Proof: see Appendix.

Proposition 2 *Given the VAR system (3), assume for simplicity that $q = 1$ (u_t is a scalar),*[4] *so that $p = 3$ and $a_y = (a_{y,1}, a_{y,2}, ..., a_{y,pk})'$, $a_w = (a_{w,1}, a_{w,2}, ..., a_{w,pk})'$ and $a_u = vec(A_u) = (a_{u,1}, a_{u,2}, ..., a_{u,pk})'$. Under the conditions of Proposition 1, the restrictions (7) take the form*

$$a_{w,1} = \frac{a_{y,i} - \gamma(a_{y,1}a_{y,i} + a_{y,p}a_{u,i} + a_{y,p+i}) - \delta}{\gamma a_{y,2} + \kappa} \quad , \quad i = 1 \tag{8}$$

$$a_{w,i} = \frac{a_{y,i} - \gamma(a_{y,1}a_{y,i} + a_{y,p}a_{u,i} + a_{y,p+i})}{\gamma a_{y,2} + \kappa} \quad , \quad 1 < i \leq p(k-1) \tag{9}$$

$$a_{w,i} = \frac{a_{y,i} - \gamma(a_{y,1}a_{y,i} + a_{y,p}a_{u,i} + a_{y,p+i})}{\gamma a_{y,2} + \kappa} \quad , \quad p(k-1) < i \leq pk. \tag{10}$$

Proof: see Appendix.

Proposition 1 ensures that it is generally possible to compute a LR test of the FL model by estimating the unrestricted VAR, and the VAR subject to $c = (2+q)pk - [(1+q)pk + 3] = pk - 3$ (over-identifying) restrictions of the form (7), where 3 is the dimension of the vector ζ. 'One-shot' tests of the FL model can be therefore computed by estimating the sysyem over the whole sample of available observations.

When the VAR in (3) is treated as agent's PLM within the adaptive learning framework, the expectations-generating system at each time t is obtained by replacing the forecasts (4)-(5) by the quantities

$$E(\widetilde{y}_{t+j} \mid \mathcal{H}_{t-1}) = s_y' A_{\mathcal{H}_{t-1}}^{j+1} X_{t-1}^*$$
$$E(\widetilde{w}_{t+j} \mid \mathcal{H}_{t-1}) = s_w' A_{\mathcal{H}_{t-1}}^{j+1} X_{t-1}^*,$$

where the symbol $A_{\mathcal{H}_{t-1}}$ here means that the estimation of the model used to forecast is based on the information available at time $t-1$, $\mathcal{H}_{t-1}$. This also means that since agents test the validity of the FL model at each t, the following sequence of restrictions arise:

$$s_y' A_{\mathcal{H}_{t-1}}(I_n - \gamma A_{\mathcal{H}_{t-1}}) - \delta s_y' - \kappa s_w' A_{\mathcal{H}_{t-1}} = 0_{1\times pk} \quad , \quad t = T_0, T_0+1, ... \tag{11}$$

where T_0 is the 'first monitoring' time.

[4]Proposition 2 can be easily extended to the case $q > 1$.

3 Cointegrated variables

Researchers dealing with FL models often find that the variables they are modelling display high persistence. This happens even when these variables are constructed as (log) deviations from their steady states, as it is the case with the equations comprising dynamic stochastic general equilibrium models. In these circumstances, a crucial decision for the purpose of reliable inference, is to establish whether time-series are better approximated as stationary (albeit persistent) processes, or as I(1), possibly cointegrated, processes, see Johansen (2006) and Fanelli (2007b).

The cross-equation restrictions derived in Section 2 have been obtained with reference to stationary processes, or considering situations where researcher's objective is testing the VAR coefficient restrictions, ignoring the possible presence of unit roots. In this section we extend the method to the case where the empirical analysis of the FL model is addressed after that the number of unit roots in the system has been opportunely fixed.

Two parameterizations of the FL model (1) are worth discussing when variables are I(1) and cointegrated.[5] The former is given by

$$\Delta\widetilde{y}_t = \psi E_t \Delta\widetilde{y}_{t+1} - \omega(\widetilde{y}_t - \phi\widetilde{w}_t) + \rho_t^* \tag{12}$$

where

$$\begin{aligned}
\psi &= \gamma/\delta, &(13)\\
\omega &= (1-\delta-\gamma)/\delta, &(14)\\
\phi &= \kappa/(1-\delta-\gamma) &(15)
\end{aligned}$$

and $\rho_t^* = \rho_t/\delta$, and holds upon the condition that $\delta + \gamma \neq 1$. The latter parameterization is based on the restriction $\delta + \gamma = 1$, and is obtained by manipulating equation (1) in the form

$$\Delta\widetilde{y}_t = \frac{1-\delta}{\delta} E_t \Delta\widetilde{y}_{t+1} + \frac{\kappa}{\delta}\widetilde{w}_t + \rho_t^* \tag{16}$$

which emphasizes that $\widetilde{w}_t$ behaves as the driving variable of $\Delta\widetilde{y}_t$. Both formulations are consistent with the case where the variables of the FL model are I(1) and possibly cointegrated; however, whereas (12) is based on the hypothesis that $\widetilde{y}_t$ and $\widetilde{w}_t$ are driven by a common stochastic trend, (16) requires that $\widetilde{w}_t$ corresponds to a (trivial) cointegrating relation.

Consider the Vector Equilibrium Correction (VEqC) representation of the I(1) 'detrended' process $\widetilde{X}_t = X_t - d_t = (\widetilde{y}_t, \widetilde{w}_t, \widetilde{u}_t')'$

$$\Delta\widetilde{X}_t = \alpha\beta'\widetilde{X}_{t-1} + \sum_{j=1}^{k-1}\Gamma_i \Delta\widetilde{X}_{t-j} + \varepsilon_t \tag{17}$$

[5] If the variables entering the FL model are I(1) and not cointegrated, the procedure described for stationary variables can be applied to a version in first differences of both the FL model and the VAR system.

where $\alpha\beta' = -(I_p - \sum_{i=1}^{k} A_i)$, α and β are $p \times r$ full rank matrices, and $\Gamma_i = -\sum_{j=i+1}^{k} A_i$, $i = 1, ..., k-1$, see Johansen (1996) for details. If y_t and w_t are cointegrated and conform to (12) and there are no additional cointegration relations in the system,

$$\beta' = (1, -\phi, 0_{1\times q}) \tag{18}$$

implying that $r = 1$, and that there are q over-identifying restrictions on β. On the other hand, if (16) is the 'correct' FL model, then

$$\beta' = (0, 1, 0_{1\times q}) \tag{19}$$

implies that there are $q+1$ over-identifying restrictions on β. If the cointegration matrix of the VEqC does not match neither the structure (18) nor (19), agent's PLM is at odds with the long-run empirical implications of the FL model.

Suppose that β complies with (18), so that $o_t = \beta'\widetilde{X}_t = \widetilde{y}_t - \phi\widetilde{w}_t$ $(r = 1)$ measures deviations from equilibrium. Define the $p \times 1$ process $W_t = (\widetilde{X}_t'\beta, \Delta\widetilde{X}_t'\tau)'$, where τ is a $p \times (p-r)$ selection matrix such that $\det(\tau'\beta_\perp) \neq 0$; Paruolo (2003), Theorem 2, shows that the VEqC (17) can be written in terms of a VAR process for $W_t = (\widetilde{X}_t'\beta, \Delta\widetilde{X}_t'\tau)'$ without loss of information, *i.e.*[6]

$$B(L)W_t = \varepsilon_t^0 \tag{20}$$

where $\varepsilon_t^0 = (\varepsilon_t'\beta, \varepsilon_t'\tau)'$ is a MDS with respect to $\mathcal{H}_t$, and

$$B(L) = I_p - B_1L - \cdots - B_k^c L^k \quad , \quad B_k^c = (B_{1k}, 0_{p\times(p-r)}),$$

with the coefficients in B_i, $i = 1, ...k$ which depend opportunely on VEqC coefficients α, Γ_1,..., Γ_{k-1}. The VAR (20) can be cast in companion form, and agents' j-step ahead conditional forecast of Δy_t and o_t can be computed as

$$E(\Delta\widetilde{y}_{t+j} \mid \mathcal{H}_t) = s_y' B^j W_t^*$$
$$E(o_{t+j} \mid \mathcal{H}_t) = s_o' B^j W_t^*$$

where s_y and s_o are selection vectors, B is the $pk \times pk$ companion matrix and $W_t^* = (W_t', W_{t-1}', ..., W_{t-k+1}')'$ is the $n \times 1$ state vector.

By the same route of Section 2, the cross-equation restrictions between the VAR (20) and the FL model (12) can be written as

$$s_y' B(I_{pk} - \psi B) + \omega s_o' B = 0_{1\times pk} \tag{21}$$

hence their recursive counterpart is given by

$$s_y' B_{\mathcal{H}_{t-1}}(I_{pk} - \psi B_{\mathcal{H}_{t-1}}) + \omega s_o' B_{\mathcal{H}_{t-1}} = 0_{1\times pk} \quad , \quad t = T_0, T_0+1, ... \tag{22}$$

[6]Johansen (1996) shows that the VEqC (17) can be written in terms of the process $W_t = (X_t'\beta, \Delta X_t'\beta_\perp)'$; Paruolo (2003) extends this result to a more general set-up.

where T_0 is the first monitoring time, and $B_{\mathcal{H}_t} \equiv B_t$ denotes the companion matrix of (20), whose parameters have to be estimated on the basis of the information available at time t.

Proposition 1, opportunely adapted, can be still applied.[7] The problem in this case, however, is that when considering the recursive tests, the cointegration matrix β in $W_t = (\widetilde{X}_t'\beta, \Delta\widetilde{X}_t'\tau)'$, must be known a priori, or must be estimated super-consistently on the basis of information $\mathcal{H}_{T_0-1}$.[8] Clearly, the analysis based on W_t requires that there are no structural breaks characterizing β from T_0 to $T^{\max}$.

A remarkable feature of the cointegrated FL model (12) is that due to the mapping (15), one of the structural parameters is directly related to β. For example, given $\beta' = (1, -\phi, 0_{1\times q})$ and γ and δ, the value of κ can be retrieved from (15).

4 Simulation-based tests

In this section we consider the problem of testing the cross-equation restrictions implied by the FL model through simulation-based techniques. For easy of reference, we shall refer to the system (3) and the restrictions (6) and their recursive counterpart (11); concepts, however, can be natually extended to the system (20) and the implied restrictions (21) and their recoursive counterpart (22).

In principle, for a fixed t, one can estimate the unrestricted VAR, and the VAR subject to the cross-equation restrictions, and compute LR statistics of the form

$$LR_t = 2(\log L(\widehat{\theta}_t) - \log L(\overline{\theta}_t)) \tag{23}$$

where $L(\cdot)$ is the likelihood function, $\widehat{\theta}_t$ is the vector containing the maximum likelihood (ML) estimates of the unrestricted VAR parameters based on the information $\mathcal{H}_t$, and $\overline{\theta}_t$ is the vector containing the constrained ML (CML) estimates of the system based on the information $\mathcal{H}_t$. The unrestricted estimates are given by $\widehat{\theta}_t = (\widehat{a}_t', \widehat{\pi}_t', \widehat{\xi}_t')'$, where $\widehat{a}_t = (\widehat{a}_{yt}', \widehat{a}_{wt}', \widehat{a}_{ut}')'$, $\widehat{\pi}_t = (\widehat{\mu}_t', vec(\widehat{\Theta}_t)')'$ and $\widehat{\xi}_t = vech(\widehat{\Sigma}_{\varepsilon t})$, whereas the constrained ones are given by $\overline{\theta}_t = (\overline{v}_t', \overline{\pi}_t', \overline{\xi}_t')$, $\overline{v}_t = (\overline{a}_{yt}', \overline{a}_{ut}', \overline{\zeta}_t')'$, $\overline{\pi}_t = (\overline{\mu}_t', vec(\overline{\Theta}_t)')'$, and $\overline{\xi}_t = vech(\overline{\Sigma}_{\varepsilon t})$. With 'one-shot' tests, the time index t in (23) is equal to $t = T^{\max}$; in the adpative learning framework t ranges from the first monitoring time T_0 onwards.

In the presence of stationary processes, comparing each LR_t with the $\chi^2_{1-\eta}(c)$ quantile, where according to Proposition 1, $c = pk - \dim(\zeta) = pk - 3$ is the number of over-identifying restrictions and η the nominal level of the test, will

[7] In this case, the number of over-identifying restrictions is given by $c = pk - (p - r) - 2$, where $2 = \dim(\psi, \omega)$, and r is the cointegration rank of the system, see Fanelli (2007b).

[8] Another possibility is represented by the recursive updating of the estimator of β. This issues is the subject of future research.

imply a rejection frequency for the null which approaches the nominal level for each fixed but large value of t, but which will result in severe size distortions if the quantile $\chi^2_{1-\eta}(c)$ is kept fixed through the entire sequence $t = T_0, T_0+1, ..., T^{\max}$. Indeed, by the law of iterated logarithms it follows that

$$\lim_{t\to\infty} P[LR_t \geq \chi^2_{1-\eta}(c)] = 1$$

also under the null, see *e.g.* Inoue and Rossi (2005).

Thus if 'one-shot' LR tests for the cross-equation restrictions (6) may exhibit size distorsions due to finite sample issues (*i.e.* when $T^{\max}$ is not large enought), their sequential counterparts are inherently size distorted if the fixed quantile $\chi^2_{1-\eta}(c)$ is used for the increasing values of t. Inoue and Rossi (2005) propose a general asymptotic theory to account for these situations, however, in our set-up asymptotic critical values might not suited at the beginning of agents' learning process, *i.e.* when the values of $t = T_0$, $T_0+1, ...$ are relatively small. Also in this case finite-sample simulated-based inference, along the lines of Dufour (2006) and Dufour and Jouini (2006), seems a reasonable solution to control the level of the test.

The algorithm for computing the LMC p-values associated with each LR_t statistics in (23) can be described as follows:

1. compute, for fixed k and t, the test statistic LR_t based on the observed data, $LR_t = LR_t^s$: this requires the unrestricted and constrained estimation of the VAR(k) in equation (3) with data until time t;

2. generate the sequence of M iid vectors ε_1^1, ε_2^2, ..., ε_t^M drawn from the Gaussian distribution $N(0, \overline{\Sigma}_{\varepsilon t})$, and given the CML estimates $\overline{\theta}_t = (\overline{v}'_t, \overline{\pi}'_t, \overline{\xi}'_t)$ obtained in the step 1, construct the M pseudo-samples of lenght t:

$$\overline{A}(L)\ X_j^m = \overline{\mu}_t + \overline{\Theta}_t D_j + \varepsilon_j^m \ , \ \ m = 1, 2, ..., M \ , \ \ j = 1, 2, ..., t$$

   where the elements of the matrices in $\overline{A}(L)$ depend on $\overline{\theta}_t$ (in particular on $\overline{v}_t$), and $X_0^m = X_0$, $X_{-1}^m = X_{-1}$, ... and $X_{-1+k}^m = X_{-1+k}$ are fixed at the initial sample values;

3. for each of the M simulated pseudo-samples of lenght t, estimate the unrestricted VAR(k) and the VAR(k) subject to the restrictions (7), and compute M independent LR statistics LR_t^m, $m = 1, 2, ..., M$;

4. compute the LMC p-value associated with LR_t^s as

$$\widehat{p}(LR_t^s, \overline{\theta}_t) = \frac{M \cdot\ \widehat{G}_M(LR_t^s) + 1}{M+1} \ \ , \ \ \widehat{G}_M(LR_t^s) = \frac{1}{M}\sum_{m=1}^{M} \mathbb{I}\left\{LR_t^m - LR_t^s\right\}$$

where the notation $\widehat{p}(LR_t^s, \overline{\theta}_t)$ reflects that the simulated p-value depends on the consistent point estimates of the restricted VAR; if $\widehat{p}(LR_t^s, \overline{\theta}_t) > \eta$, the LMC test is not significant at level η;[9]

5. for the sequential (recusive) version of the test repeat the steps 1-4 for $t = T_0 + 1, T_0 + 2, ..., T^{\max}$.

The LMC power associated with each LR_t statistics can be obtained similarly, by simply generating the pseudo-samples in the step 2 from the unconstrained ML estimates $\widehat{\theta}_t = (\widehat{a}_t', \widehat{\pi}_t', \widehat{\xi}_t')'$. The procedure described above can be viewed as a degenerate version of the maximized Monte Carlo (MMC) tests proposed in Dufour (2006): instead of maximizing a simulated p-value function over a consistent set estimator of the VAR nuisance parameters, simulations in step 2 are based on consistent point estimates.

The 'price to pay' with the LMC procedure, compared to MMC techniques, is that stronger assumptions are required to yield asymptotically valid tests under general conditions. However, albeit LMC-type tests are computer intensive, they are still feasible compared to MMC-type techniques which require the maximization of a simulated p-value function over the entire nuisance's parameter space. Moreover, the simulation experiment we discuss in Section 5 shows that the suggested LMC procedure results in LR tests displaying empirical rejection frequency substantially close to the nominal level, under the null that the cross-equation restrictions hold, and reasonable power against the hypothesis of a backward-looking model with variables generated by I(1) cointegrated processes.

CML estimates in the steps 1 and 3 can be computed by combining standard Newton-like methods with a grid search for the three structural parameters $\zeta = (\gamma, \delta, \kappa)'$ as in Fanelli (2007b). This choice, followed in Section 5 and Section 6, can be computationally demanding but has the advantage that obvious a priori bounds characterizing γ, δ and κ suggested by the theory, can be easily accounted for.[10] However, in this case a balancing is required between computation costs and the necessity of specifying a sufficiently fine grid in a way that estimation and testing procedures do not turn out to be distorted.

5 Simulation experiment

In this section we present simulation evidence on the performance of LMC LR-type tests for the FL nonlinear restrictions discussed in Section 4. We consider a

[9] Strictly speaking, $\widehat{p}(LR_t^s, \overline{\theta}_t) \leq \eta$, is not sufficient to reject the null hypothesis at level η. Indeed, as explained in Dufour and Jouini (2006), the rejection of the null should be based on the maximized p-value: $\widehat{p}_{MMC} = \sup\{\widehat{p}_M(LR_t^s, \theta), \theta \in \Omega_0\}$, where Ω_0 is the parameter space restricted by the null hypothesis, or a consistent set restricted estimator.

[10] Alternatively, suitable parameter transformations can be implemented as described in Dejong and Dave (2007), Chapter 8.

simple DGP, represented by a VAR process with $p = 2$ variables and $k = 2$ lags. Results are reported in Table 1.[11]

Data are generated from the DGP under the null that the VAR companion matrix A is subject to the nonlinear cross-equation restrictions (6) (see Proposition 1), and setting the structural parameters of the FL model at the values $\zeta^* = (0.70, 0.20, 0.15)'$.[12] In order to mimic the typical time-series persistence researchers face empirically, the eigenvalues of the (constrained) companion matrix A are specified such that the highest root is close to unity.

The nominal level of the test is 0.05. LMC LR-type tests are based on M =100 replications (pseudo-samples), while the number of trials used for evaluating rejection frequencies is 1000.[13] Since our emphasis is on the typical sample lengths available to macroeconometricians, we consider samples of length T=50 and T =100, respectively.

The first and second columns of Table 1 summarize the percentages of rejections of the FL model under the null that the cross-equation restrictions hold. The number of over-identifying restrictions in the LR tests is $c = pk - \dim(\zeta) = 1$, see Proposition 1. CML estimates are obtained by combining Newton-like methods with a grid search for ζ; the grid is constructed around the point $\zeta^* = (0.70, 0.20, 0.15)'$, using ± 0.02 as increment; $\zeta^{upper} = (0.74, 0.24, 0.19)'$ and $\zeta^{lower} = (0.66, 0.16, 0.11)'$ are taken as the upper and lower bounds of the grid, respectively.

Power in the third and forth columns of Table 1 is obtained under the alternative that the process is generated by the unrestricted VAR in (20), where $W_t = (X_t'\beta, \Delta X_t'\tau)'$ is equal to $W_t = (X_{1t} - X_{2t} - 0.2, \Delta X_{2t})'$; the cointegration matrix is hence fixed at $\beta' = (1, -1, -\mu_0)$ and $\beta' = (1, -1, 0, -\mu_0)$, respectively, where $\mu_0 = 0.2$ is a constant restricted to lie in the cointegration space (Johansen, 1996).[14] This choice of the alternative reflects the situation where the FL restrictions are not supported by the data, and the variables entering the model are cointegrated.

Rejection proportions in Table 1 are expressed in percentages. They show that when the nonlinear FL cross-equation restrictions hold, asymptotic tests based on standard critical values display non-negligible overrejection rates (for example, 0.215 instead of 0.05). This simple simulation confirms that the rejection of FL models may be partly ascribed to the use of asymptotic critical values in finite

[11]The simulation experiment has been performed through Ox 4.0. All codes, including those relative to the estimates of Section 6 are available upon request.

[12]Gaussian disturbances with covariance matrices having 1 on the diagonal and 0.5 as off-diagonal elements are used. Observe that in this experiment as well as in Section 6, we consider a parametric version of the LMC procedure (parametric bootstrap, also known as MC bootstrap); in principle, however, also a non-parametric version of these procedures might be implemented.

[13]Results very similar to those reported in Table 1 can be obtained by increasing both M and the number of trials used to evaluate rejection frequencies.

[14]The highest root of the VAR companion matrix of the W_t process is 0.67 for DGP1, and 0.90 for DGP2.

samples. In this respect, the LMC procedure seems to control the level of the test satisfactorily, even when the VAR roots are close to unity. Interestingly, the power of the LMC LR-type tests against the alternative of a backward-looking specification with I(1) cointegrated processes, appears reasonable.

6 Application to the New Keynesian Phillips curve

The NKPC can be regarded as the most prominent FL model of inflation dynamics that currently dominates the debate in moneary policy. It reads as a special case of the equation (1), where $\widetilde{y}_t$ is the inflation rate and $\widetilde{w}_t$ a measure of firms' real marginal costs. In the Calvo (1983) formulation, the structural parameters in $\zeta = (\gamma,\delta,\kappa)'$ are related to other 'deep parameters' through the mapping

$$\gamma = \frac{\varrho\vartheta}{\vartheta + \tau[1-\vartheta(1-\varrho)]} \tag{24}$$

$$\delta = \frac{\tau}{\vartheta + \tau[1-\vartheta(1-\varrho)]} \tag{25}$$

$$\kappa = \frac{(1-\tau)(1-\vartheta)(1-\vartheta\varrho)}{\vartheta + \tau[1-\vartheta(1-\varrho)]} \tag{26}$$

where ϱ is firms' discount factor, τ the fraction of forward-looking firms, and $(1-\vartheta)^{-1}$ is the average time upon which prices are kept fixed, see Galí and Gertler (1999) and Galí *et al.* (2001). Using a generalized method of moments approach, Galí *et al.* (2001) find support for the NKPC in the Euro area over period 1971-1998. Other existing studies based on model consistent expectations (Bårdsen *et al.*, 2004) and 'VAR expectations' (Fanelli, 2007b), reject the NKPC for the euro area using the same data set as Galí *et al.* (2001).

6.1 Data

We consider quarterly data for the euro area, using the last release of the Area-wide Model (AWM) data set described in Fagan *et al.* (2001). Variables cover the period 1971:1-2005:4. To measure inflation we use the GDP deflator, *i.e.* $y_t = 4 \times 100(p_t - p_{t-1}) = \pi_t$, where p_t is the log of the GDP deflator. The GDP deflator is YED in the AWM data set. Firms' average marginal costs are proxied by the wage share (log of real unit labour costs), $w_t = ws_t$. The wage share (real unit labour costs) is computes as $ws_t = 100 \times \log(WIN_t/YER_t)$, where WIN is 'Compensation to Employees' (in real terms) and YER is real GDP. A short term interest rate, $u_t \equiv i_t$ $(q_u = 1)$, is included in the system. As a proxy of i_t, we have used the short-term interest rate, which is STN in the data set.

The estimation and testing procedure under the VAR-based adaptive learning dynamics described in Section 2 and Section 4 relies on the sub-sample 1971:1-1983:4 to produce initial estimates of VAR parameters, and uses the sub-period 1984:1-2005:4 to evaluate the NKPC recursively.[15] In terms of the notation of the previous sections, T_0 =1984:1 is the first monitoring period, and $T^{\max}$ =2005:4 is the last available observation. Clearly, 'one-shot' tests are carried out over the entire period 1971:1-2005:4.

6.2 Results

We consider a VAR for $X_t = (\pi_t, ws_t, i_t)'$ with three lags ($k = 3$) and a costant (*i.e.* $A(L)d_t = \mu$, $\Theta \equiv 0, D_t \equiv 0$). The constant is restricted to lie in the cointegration space ($\mu = \alpha\mu_0$, where μ_0 is the intercept entering the cointegratign relation) when we perform the cointegration rank test, since the data do not show any linear trend in the variables. Table 2 reports the estimated roots of companion matrix, the LR cointegration trace test, and the cointegrating relations, over both the sub-period 1971:1-1983:4, and the whole sample 1971:1-2005:4. Results based on the entire sample of observation refer to the 'one-shot' analysis of the NKPC, while results based on the sub-period 1971:1-1983:4 read as agents' initial beliefs in the adaptive learning framework.

Table 2 shows that the highest estimated root of VAR companion matrix is close to unity, emphasizing that the system is highly persistent. Actually, the LR trace test for cointegration rank suggests that the hypothesis of unit roots is highly supported by the data; moreover, inflation and the wage share seem to co-move in the euro area, confirming the finding in Fanelli (2007b), based on a previous release of the AWM data set.

Although the evidence in Table 2 remarks that the system comprising euro area inflation, the wage share and the short term interest rate might be reasonably approximated as an I(1) cointegrated VAR, we recognize that the power of the cointegration rank test might be poor against the alternative of an highly persistent but stationary process in finite samples. For this reason, we take an 'agnostic' view on the issue, and consider two possible specifications of agents' expectations generating system: in the former, discussed in Subsection 6.2.2, the VAR in levels is used to investigate the NKPC along the lines of Section 2, without imposing the cointegration restrictions; in the latter, discussed in Subsection 6.2.3, the nonstationarity and cointegration restrictions of the system are fully incorportated in the econometric analysis, as outlined in Section 3.

In both approaches, the estimation of the constrained VAR is achieved by combining Newton-like methods with a grid search for the structural parame-

[15]Fanelli (2007a) provides a detailed explanation of why 1984:1 can be chosen as the first monitoring time. The results of this section, however, are robust to changes in the first monitoring time, and are available upon request.

ters of the NKPC; the next Subsection explains how the grid for the structural parameters had been constructed.

6.2.1 The grid

CML estimation is carried put by combining like-Newton methods with a grid search for $\zeta = (\gamma,\delta,\kappa)'$. This method, already used in Fanelli (2007b), is in this context feasible because the dimension of ζ is not large, and allows to consider values for the structural parameters which are a priori consistent with the Calvo set-up, see the equations (24)-(26).

The grid used in Subsection 6.2.2 is constructed as follows: starting from $\gamma = 0.05$, $\delta = 0.05$ and $\kappa = 0.05$, 0.01 is used as incremental value for γ and δ, and 0.05 is used as incremental value for κ. The following constraints and bounds are imposed: $\gamma > 0$, $\delta > 0$, $\gamma + \delta < 1$, $0 < \kappa \leq 0.4$, obtaining a grid comprising 8280 points. The restriction $\gamma + \delta = 1$ is ruled out, since preliminary computations reveal that the data do not comply with a formulation of the NKPC of the form (16), see Section 3; the upper bound for κ is motivated by theoretical considerations.

The mapping (13)-(14) and the grid for ζ are then used in Subsection 6.2.3 to construct a grid for ψ and ω when the cointegrated version of the NKPC is estimated and tested; for fixed ψ and ω and β (see in particular equation (18)), the mapping (15) provides values for κ.[16]

6.2.2 Model in levels

For $t = T_0$, $T_0 + 1$, ..., $T^{\max}$, a VAR(3) for $X_t = (\pi_t, ws_t, i_t)'$ is estimated recursively under the cross-equation restrictions (8)-(9), and unrestrictedly. The CML estimates show that, for all t, except t =1984:1, the restricted log-likelihood is maximized in correspondence of the structural parameters vector $\widehat{\zeta}_t = \zeta^* = (0.93, 0.05, 0.40)'$.[17] In terms of the Calvo parameterization provided by equations (24)-(26), the counterparts of these estimates are $\varrho^* = 0.98$ (discount factor), $\tau^* = 0.028$ (fraction of backward-looking firms) and $(1 - \vartheta^*)^{-1} \approx 2$ (average number of quarters over which prices are kept fixed).

The upper-panel of Figure 1 plots the entire sequence of LR statistics; the graph also reports the 95% quantile taken from a $\chi^2_{0.95}(6)$. Using this quantile,

[16]Preliminary estimates have shown that, in practice, computation time can be firmly reduced without affecting estimation and testing results, by considering only a limited number of grid points (50 points). In particular, the region of (structural) parameter space where the magnitude of γ (forward-looking parameter) dominates that of δ (backward-looking parameter) seems to be preferred in terms of likelihood.

[17]For t =1984:1 we obtain $\widehat{\zeta}_t = (0.87, 0.11, 0.40)'$. Given that this estimate does not differ too much from the (stable) one obtained over the period 1984:2-2005:4, in this section the analysis will be developed by assuming that $\widehat{\zeta}_t = (0.93, 0.05, 0.40)' = \zeta^*$, all t.

the FL model of inflation dynamics is sharply rejected for all $t = T_0$, $T_0 + 1$, ..., $T^{\max}$. Hence for $t = T^{\max}$, the 'one-shot' test rejects the NKPC for the euro area.

The lower-panel of Figure 1 plots the sequence of simulated p-values associated with each LR statistic, computed through the LMC procedure described in the steps 1-5 of Section 4. More precisely, for each $t = T_0, T_0{+}1, ..., T^{\max}$ we generated $M = 1000$ pseudo-samples drawn from a VAR(3) with coefficients fixed at the CML sample estimates $\overline{\theta}_t = (\overline{a}'_{yt}, \overline{a}'_{ut}, \overline{\zeta}'_t, \overline{\mu}'_t, \overline{\xi}'_t)$, $\overline{\xi}_t = vech(\overline{\Sigma}_{\varepsilon t})$. The lower-panel of Figure 1 also plots the simulated rejection frequencies (power) associated with each LR statistic; here, the pseudo-samples are generated from a VAR(3) with coefficients fixed at the unrestricted sample estimates $\widehat{\theta}_t = (\widehat{a}'_{yt}, \widehat{a}'_{wt}, \widehat{a}'_{ut}, \widehat{\mu}'_t, \widehat{\xi}'_t)'$, $\widehat{\xi}_t = vech(\widehat{\Sigma}_{\varepsilon t})$. The graph shows that using simulated p-values, each test in the sequence is not significant at the 5% level.

Estimation suggests that agents form their beliefs consistently with the theoretical restrictions implied by the NKPC over the period 1984:1-2005:4. Similarly, the last LR test in Figure 1 remarks that when $t = T^{\max}$, the LMC 'one-shot' test of the NKPC does not reject the model, albeit marginally, at the 0.05 level (simulated p-value equal to 0.053). This result differs from those obtained using the standard $\chi^2_{1-\eta}$ quantile.

6.2.3 Cointegrated model

Table 2 shows that the system $X_t = (\pi_t, ws_t, i_t)'$ can be approximated as an I(1) cointegrated processes in the euro area, where in particular π_t and ws_t comply with (18), and the short term interest rate is weakly exogenous. Hereafter $\widehat{\beta}' = (1, -\widehat{\phi}, 0, \widehat{\mu}_0)$ with $\widehat{\phi} = 0.5$ and $\widehat{\mu}_0 = 1.88$ (μ_0 is the restricted constant), will be taken as the estimated cointegration vector.[18]

In this framework, the empirical analysis of the NKPC is based on the formulation (12), and the cointegrated VAR for $X_t = (\pi_t, ws_t, i_t)'$ is re-specified in the form (20), with $\widehat{\beta}' X_t = o_t = \widetilde{\pi}_t - 0.5\widetilde{ws}_t + 1.88$, and where the 3×2 matrix τ selects $\Delta\pi_t$ and Δi_t from ΔX_t ($\det(\tau'\widehat{\beta}_\perp) \neq 0$). The expectations generating system is thus represented by a stable VAR(3) for the vector $W_t = (o_t, \Delta\pi_t, \Delta i_t)'$.

For $t = T_0$, $T_0 + 1$, ..., $T^{\max}$, the system is estimated recursively under the cross-equation restrictions (22), and unrestrictedly. The resulting CML estimates are given by $(\widehat{\psi}_t, \widehat{\omega}_t)' = (\psi^*, \omega^*)' = (18.6, 0.4)'$ for all t. According to (13)-(14) and (15), the corresponding structural parameters are $\widehat{\zeta}_t = \zeta^{**} = (0.93, 0.05, 0.01)'$, for all t. Note, in particular, that the value $\kappa^{**} = (1 - \gamma^{**} - \delta^{**})\widehat{\phi} = 0.01$ is obtained by inverting the relation (15). Thus, it follows that two of the three estimated structural parameters coincide with the estimates obtained through the VAR in levels; in terms of the Calvo parameterization, the vector ζ^{**} implies $\varrho^{**} = 0.98$, $\tau^{**} = 0.048$ and $(1 - \vartheta^{**})^{-1} \approx 11$. In this context, compared to

[18] Note that a similar results is obtained in Fanelli (2007b) using a previous release of the AWM data set.

the estimates obtained in Subsection 6.2.2, the probability that firms are able to change their prices across periods falls from $(1-\vartheta^{*}) \approx 0.5$ to $(1-\vartheta^{**}) \approx 0.09$.

The upper-panel of Figure 2 plots the entire sequence of LR statistics; the graph also reports the 95% quantile taken from $\chi^2_{0.95}(5)$.[19] Using this quantile, the FL model of inflation dynamics is rejected from $t =$1999:2 onwards. Hence, also in this case, for $t = T^{\max}$ the 'one-shot' test rejects the (cointegrated) NKPC as in Fanelli (2007b).

The lower-panel of Figure 2 plots the sequence of simulated p-values and power associated with each LR statistic, computed through the LMC procedure. It turns out that also the cointegrated version of the NKPC is supported by the data, with reasonable power during the period 1984:1-2005:4. As with the VAR in levels, when $t = T^{\max}$, the LMC 'one-shot' test for the cross-equation restrictions rejects the NKPC only marginally at the 0.05 level (simulated p-value equal to 0.045).

To sum up, when the cointegration restrictions are incorporated in the agents' expectations generatign system, the NKPC for the euro area is still supported likewise the model in levels; however, in this case the magnitude of the estimated κ parameter, which is directly related to the pass-through from marginal costs to inflation, is remarkably lower, implying an higher degree of price stickiness.

7 Conclusions

In this paper we have proposed a method which allows to investigate the finite sample performance of tests for the nonlinear cross-equation restrictions that arise when FL models are tested against VAR systems. We consider two types of LR tests: (i) 'one-shot' tests, where estimation involves the entire sample of observations, and which are based on the comparison between the unrestricted and constrained VAR likelihoods; (ii) sequences of tests obtained under the assumption that 'boundedly rational' agents use the VAR recursively to update their beliefs through the perpetual assessment of the FL model, also based on the comparison of the unrestricted and constrained system likelihoods.

A simple simulation experiment highlights that the use of asymptotic critical values taken from the chi-squared distribution may induce the researcher to falsely reject FL models in finite samples. This result helps to explain why, beyond the well-established idea that FL models are inherently misspecified representations of the data, they receive poor (or marginal) empirical support in the literature. We show that LMC LR-type statistics allow a substantial control of the level of the test, even when the persistence of the modelled time-series is high; in addition, LMC tests exhibit reasonable power against the alternative of a backward-looking specification where the variables are generated by I(1) cointegrated processes. Hence, at first glance, simulation-based inference can justify the practise of testing

[19]See footnote 6 for details.

FL models against VARs with variables in levels, without pre-testing for unit roots and cointegration. Nevertheless, our application to the NKPC for the euro area over the period 1971-2005 shows that this is not always the case.

More precisely, our approach shows that, whereas tests based on asymptotic critical values taken from the chi-squared distribution reject the model, both 'one-shot' and recursive LMC LR-type tests tend to support (albeit marginally for the 'one-shot' tests) the nonlinear constraints that the NKPC imposes on the VAR. Moreover, the magnitude of the forward-looking parameter of the NKPC dominates that of the backward-looking parameter, confirming a well-established finding of the econometric adaptive learning literature, where it is often argued that learning mechanisms can replace those devices, such as indexation, contracts, rules of thumb, etc., which are included to capture inflation persistence.

The euro area NKPC seems to be supported, using simulation-based inference, irrespective of whether the system is treated as a stationary, or as an I(1) cointegrated system, where inflation and the wage share co-move over time. More importantly, however, compared to the results obtained using the VAR in levels, the imposition of cointegration restrictions changes the estimated magnitude of one of the structural parameters, with remarkable consequences on the implied degree of price stickiness.

In this paper we take an 'agnostic' position on whether agents' expectations generating system is better approximated as a stationary (but highly persistent), or non-stationary cointegrated VAR; overall results, however, suggest that the choice between these two options plays an important role in the empirical assessment process.

Appendix

Proof of Proposition 2

Given the definition of s_y, s_w and s_u, observe that $A' = (a_y, a_w, A_u, \Xi)$, where Ξ is the sub-matrix of the companion form matrix containing 0 and 1 only, hence $vec(A') = (a', vec(\Xi)')'$, where $a = (a'_y, a'_w, a'_u)'$ and $a_u = vec(A_u)$. Define the function

$$f(a_w, a_y, a_u, \zeta) = f(a_w, v) = (I_n - \gamma A')a_y - \delta s_y - \kappa\ a_w$$

where $f : \mathbb{S} \rightarrow \mathbb{R}^n$, $n = pk$, $\mathbb{S}$ is an open set in $\mathbb{R}^{(q+2)n+3}$, $3 = \dim(\zeta)$, and $v = (a'_y, a'_u, \zeta')'$. Under the null that the restrictions (6) hold, one has

$$f(a_w^0, v^0) = 0_{n\times 1} \tag{27}$$

where a_w^0 and v^0 denote the 'true' parameter values of a_w and v, respectively, and (a_w^0, v^0) is an interior point of $\mathbb{S}$. The function f is twice differentiable at (a_w^0, v^0).

The $n \times n$ Jacobian $J(a_w, v) = \partial f(a_w, v)/\partial a'_w$ is given by

$$\frac{\partial f(a_w, v)}{\partial a'_w} = \underbrace{\frac{\partial \left\{ (I_n - \gamma A')a_y \right\}}{\partial a'_w}}_{I_n} - \kappa \underbrace{\frac{\partial a_w}{\partial a'_w}}_{I_n}$$

$$= -\gamma \frac{\partial \left\{ A' a_y \right\}}{\partial a'_w} - \kappa I_n = -\gamma \frac{\partial vec(A' a_y)}{\partial a'_w} - \kappa I_n$$

$$= -\gamma (a'_y \otimes I_n) \frac{\partial vec(A')}{\partial a'_w} - \kappa I_n = -\gamma (a'_y \otimes I_n) \begin{bmatrix} 0_{n \times n} \\ I_n \\ 0_{[n(n-2)] \times n} \end{bmatrix} - \kappa I_n$$

$$= -\gamma (a_{y,1} I_n, a_{y,2} I_n, ..., a_{y,n} I_n) \begin{bmatrix} 0_{n \times n} \\ I_n \\ 0_{[n(n-2)] \times n} \end{bmatrix} - \kappa I_n$$

$$= -(\gamma a_{y,2} + \kappa) I_n. \tag{28}$$

It turns out that $J(a_w, v)$ is non-singular at (a^0_w, v^0) iff $a^0_{y,2} \neq -(\kappa^0/\gamma^0)$. Hence, if $a^0_{y2} \neq -(\kappa^0/\gamma^0)$, by the implicit function theorem there exists an open set $\mathbb{T}$ in $\mathbb{R}^{(q+1)n+3}$ containing v^0, and a unique differentiable function $g : \mathbb{T} \to \mathbb{R}^n$, such that

$$a^0_w = g(v^0) \tag{29}$$

and $f(g(v), v) = 0_{n \times 1}$ for all $v \in \mathbb{T}$.

The number of free parameters of the VAR under the null is $(1+q)n+3$, hence the number of over-identifying restrictions is $c = (2+q)n - [(1+q)n+3] = n-3$. This implies that $pk \geq 4$ for the restrictions to be binding, *i.e.* for $c \geq 1$. This proves *(i)* and *(ii)*.

Given the function (29), the mapping between the parameters of the unconstrained VAR and the parameters of the restricted VAR can be written in explicit form as

$$a = h(v)$$

where, given the partion $a = (a'_y, a'_w, a'_u)'$, the function $h(\cdot)$ has the following structure:

$$\begin{aligned} a_y &= h_y(v) = a_y \\ a_w &= h_w(v) \equiv g(v) \\ a_u &= h_u(v) = a_u. \end{aligned}$$

Let $logL$ be the log-likelihood of the VAR. Using the chain rule of derivatives the score associated with v is given by

$$\frac{\partial logL}{\partial v'} = \underset{1 \times [(2+q)n]}{\frac{\partial logL}{\partial a'}} \times \underset{[(2+q)n] \times [(1+q)n+3]}{\frac{\partial h(v)}{\partial v'}}$$

where $(\partial logL/\partial a)$ is the score associated with the parameters of the unrestricted system, and the Jacobian is defined as

$$\frac{\partial h(v)}{\partial v'} = D(v) = \left[\begin{array}{ccc} I_n & 0_{n\times nq} & 0_{n\times 3} \\ D_{g,a_y} & D_{g,a_u} & D_{g,\zeta} \\ 0_{nq\times nq} & I_{nq} & 0_{qn\times 3} \end{array} \right] \equiv [D_1(v)\ ,\ D_2(v)\ ,\ D_3(v)] \qquad (30)$$

where $D_{g,a_y} = \partial g(v)/\partial a_y'$ is $[(1+q)n+3]\times n$, $D_{g,a_u} = \partial g(v)/\partial a_u'$ is $[(1+q)n+3]\times nq$, and $D_{g,\zeta} = \partial h_w/\partial \zeta'$ is $[(1+q)n+3]\times 3$. The matrix $D(v)$ in (30) has full column rank $(1+q)n+3$ at $v=v^0$ if and only if $D_3(v^0)$ has full column rank 3, and the columns of the sub-matrices $D_1(v^0)$, $D_2(v^0)$, $D_3(v^0)$ are linearly independent. It can be noticed that by construction the columns of $D_1(v)$ can not be expressed as linear combination of the columns of $D_2(v)$ and $D_3(v)$, whereas the columns of $D_2(v)$ can not be expressed as linear combinations of the columns of $D_1(v)$ and $D_3(v)$. Similarly, the columns of $D_3(v)$ can not be obtained as linear combinations of the columns of $D_1(v)$ and $D_2(v)$. Finally, observe that $D_3(v)$ has full column rank 3 at $v=v^0$ iff $D_{g,\zeta}(v^0) = \partial g(v^0)/\partial \zeta'$ has column rank 3. More precisely,

$$D_{g,\zeta} = \frac{\partial g(v)}{\partial(\gamma,\delta,\kappa)'} = \left[\underset{n\times 1}{\frac{\partial g(v)}{\partial \gamma}}, \underset{n\times 1}{\frac{\partial g(v)}{\partial \delta}}, \underset{n\times 1}{\frac{\partial g(v)}{\partial \kappa}} \right] \qquad (31)$$

where the $g(\cdot)$ function in (29) is such that $\partial g(v)/\partial\gamma = -(\gamma a_{y,2}+\kappa)^{-1}[r(v) - a_{y,2}g(v)]$ with $r(\cdot)$ continuous function, $r:\mathbb{R}^{(1+q)n+3}\to\mathbb{R}^n$, $\partial g(v)/\partial\delta = -(\gamma a_{y,2}+\kappa)^{-1}i_1$, where i_1 is a $n\times 1$ vector with 1 as its first element and zero elsewhere, and $\partial g(v)/\partial\kappa = -(\gamma a_{y,2}+\kappa)^{-1}g(v)$. Moreover, $g(v^0)\neq i_1$, $r(v^0)\neq i_1$, $g(v^0)\neq r(v^0)$ and $r(v^0) - a_{y,2}g(v^0)\neq \lambda i_1$ for each scalar λ. Accordingly, each column of the matrix (31) evaluated at $v=v^0$ is linearly independent on the other two columns. It turns out that $D_{g,\zeta}(v^0)$ has rank 3, implying that the Jacobian $D(v^0)$ in (30) has rank $(1+q)n+3$.

Under standard regularity conditions, the $[(1+q)n+3]\times[(1+q)n+3]$ left-upper block of the information matrix of the restricted VAR is given by

$$R(v) = E\left[\frac{\partial logL}{\partial v}\times\frac{\partial logL}{\partial v'}\right] = D(v)'R(a)D(v),$$

where $R(a) = E\left[\frac{\partial logL}{\partial a}\times\frac{\partial logL}{\partial a'}\right]$ is the left-upper block of the information matrix of the unrestricted VAR, which is non-singular. $R(v^0)$ is non-singular as $D(v^0)$ has rank $(1+q)n+3$. This proves *(ii)* and completes the proof.■

Proof of Proposition 2

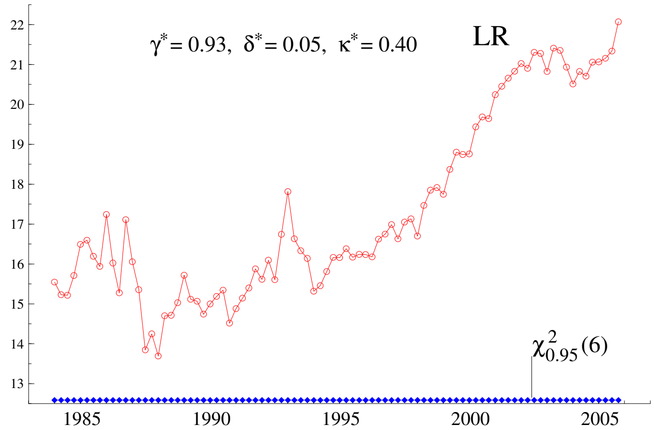
Given $p=3$, and trasposing both sides of (6) gives rise to the following system of equations

$$
\begin{aligned}
\gamma(a_{y,1}a_{y,1} + a_{y,2}a_{w,1} + a_{y,3}a_{u,1} + a_{y,p+1}) - a_{y,1} + \delta + \kappa a_{w,1} &= 0 \\
\gamma(a_{y,1}a_{y,i} + a_{y,i}a_{w,i} + a_{y,3}a_{u,i} + a_{y,p+i}) - a_{y,p+i} + \kappa a_{w,i} &= 0 \quad , \quad i = 2, ..., pk
\end{aligned}
$$

where $a_{y,p+i} = 0$ if $i > p(k-1)$. The structure of the system is such that each $a_{w,i}$, $i = 1, ..., pk$ can be expressed as unique function of the $2n+3$ parameters $a_y = (a_{y,1}, a_{y,2}, ..., a_{y,pk})'$, $a_u = (a_{u,1}, a_{u,2}, ..., a_{u,pk})'$ and $\zeta = (\gamma,\delta,\kappa)'$, consistently with equation (7). The solution of the system with respect to each element of $a_w = (a_{w,1}, a_{w,2}, ..., a_{w,pk})'$ amounts to the equations (8)-(10) reported in the text. This completes the proof. ■

Figures

Figure 1: Upper panel — sequence of LR statistics for the restrictions implied by the NKPC under learning dynamics (VAR in levels), with corresponding estimates of structural parameters $\zeta = (\gamma, \delta, \kappa)'$, and $\chi^2_{0.95}(6)$ quantile; Lower panel — LMC simulated p-values and 0.95 powers.



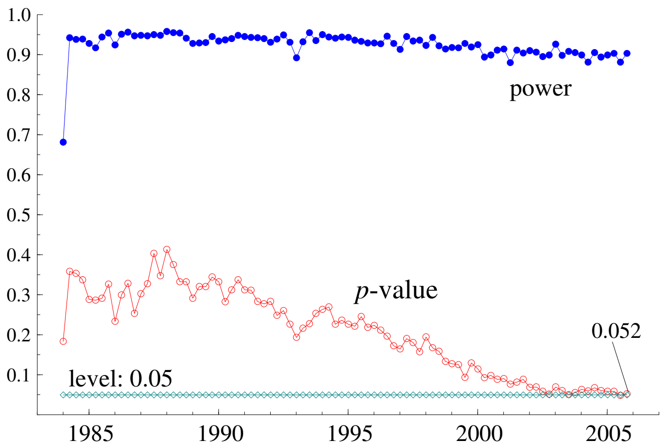
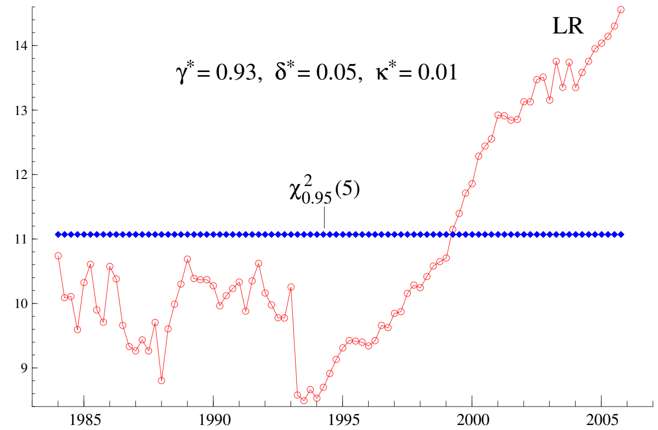
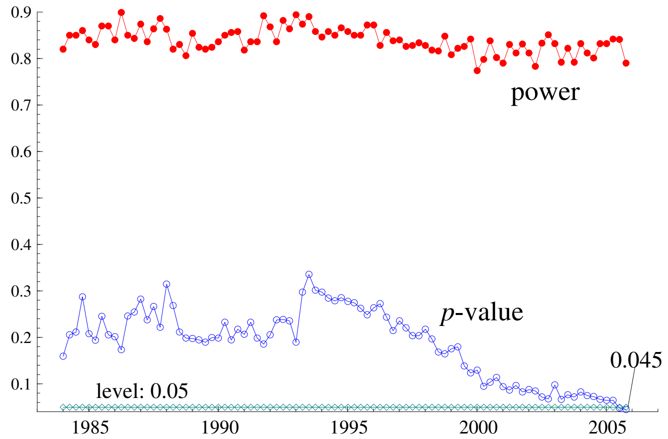


Figure 2: Upper panel — sequence of LR statistics for the restrictions implied by the cointegrated NKPC under learning dynamics, with corresponding estimates of structural parameters $\zeta = (\gamma, \delta, \kappa)'$, and $\chi^2_{0.95}(6)$ quantile; Lower panel — LMC simulated p-values and 0.95 powers.





Tables

Table 1: Rejection frequencies (percentages) of asymptotic and LMC LR-type tests for the cross-equation restrictions implied by the FL model (1) with structural parameters $\zeta = (\gamma, \delta, \kappa)' = (0.70, 0.20, 0.15)'$ on the VAR.

DGP: bivariate VAR(2) ($p = 2$, $k = 2$)				
Highest eigenvalue of (restricted) companion matrix A: 0.989				
	Level		*Power*	
	ASY_{LR}	LMC_{LR}	ASY_{LR}	LMC_{LR}
T=50	27.9	3.5	77.7	41.5
T=100	21.5	2.9	97.6	83.7

Note: the nominal level of the tests is η=0.05; ASY_{LR} stays for the asymptotic (chi-squared distributed) test based on the LR statistic, while LMC_{LR} is the corresponding LMC test. Power is obtained under the alternative of variables generated by I(1) cointegrated processes, see Section 5 for details. Rejection frequencies are based on 1000 replications and simulated p-values and powers are computed using M=100, see Section 4.

Table 2: LR trace test for cointegration rank and estimated cointegrating relation and adjustment coefficients.

VAR(3): $X_t = (\pi_t, ws_t, i_t)'$		1971:1- 1983:4 *1971:1-2005:4*
Highest eigenvalues of companion matrix:		0.92± 0.066i *0.977± 0.027i*
	Cointegration rank test	
$H_0 : r \leq j$	Trace	p-value
$j = 0$	43.74	0.004
	53.18	*0.000*
$j = 1$	13.66	0.321
	14.29	*0.276*
$j = 2$	4.79	0.318
	1.64	*0.839*

Estimated cointegrating relation and adjustment coefficients

$$\hat{\beta}' X_t = o_t = \pi_t - \underset{(0.089)}{0.491} ws_t + \underset{(0.357)}{1.88}, \quad \alpha = (\underset{(0.20)}{-0.54}, \underset{(0.22)}{0.05}, 0)'$$

$$*\hat{\beta}' X_t = o_t = \pi_t - \underset{0.043}{0.515} ws_t + \underset{0.170}{1.98}, \quad \alpha = (\underset{0.08}{-0.25}, \underset{0.03}{0.15}, 0)'*$$

Numbers by asterisks, * *, are computed on the entire sample 1971:1-2005:4. Standard errors in parentheses, p-values in brackets.

References

Baillie, R. T. (1989), "Econometric Tests of Rationality and Market Eefficiency", *Econometric Reviews*, 8, pp. 151–186.

Bårdsen, G., E. S. Jansen and R. Nymoen (2004), "Econometric Evaluation of the New Keynesian Phillips Curve", *Oxford Bulletin of Economics and Statistics*, 66 (supplement), pp. 671–685.

Bekaert, G. and R. Hodrick (2001), "Expectations Hypotheses Tests", *Journal of Finance*, 56, pp. 1357–1394.

Binder, M. and H. M. Pesaran (1995), *Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results*, in M.H. Pesaran and M. Wickens (eds.), *Handbook of Applied Econometrics*, Oxford: Blackwell.

Branch, W. A. and G. W. Evans (2006), "A Simple Rrecursive Forecasting Model", *Economic Letters*, 91, pp. 158–166.

BRAYTON, F., M. EILEEN, T. P. REIFSCHNEIDER AND J. WILLIAMS (1997), "The Role of Expectations in the FBR/US Macro-Economic Model", *Federal Reserve Bulletin*, 83, pp. 227–245.

CALVO, G. A. (1983), "Staggered Contracts in a Utility-Maximizing Framework", *Journal of Monetary Economics*, 12, pp. 383–398.

CAMPBELL, J. Y. AND R. J. SHILLER (1987), "Cointegration and Tests of Present-Value Models", *Journal of Political Economy*, 95, pp. 1052–1088.

CANOVA, F. (2007), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton.

DEJONG, D. N. AND C. DAVE (2007), *Structural Microeconometrics*, Princeton University Press, Princeton.

DUFOUR, J.-M. (2006), "Monte Carlo Tests with Nuisance Parameters: A General Approach to Finite-Sample Inference and Nonstandard Asymptotics", *Journal of Econometrics*, 133, pp. 443–477.

DUFOUR, J.-M. AND T. JOUINI (2006), "Finite-Sample Simulation-Based Inference in VAR Models with Application to Granger Causality Testing", *Journal of Econometrics*, 135, pp. 229–254.

EVANS, G. W. AND S. HONKAPOHJA (1999), *Learning Dynamics*, in Handbook of Macroeconomics 1A, Chap. 7.

EVANS, G. W. AND S. HONKAPOHJA (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton.

FAGAN, G., G. HENRY AND R. MESTRE (2001), "An Area-Wide Model (awm) for the Euro Area", *Working Paper, European Central Bank*, No. 42.

FANELLI, L. (2007a), "Evaluating the New Keynesian Phillips Curve Under VAR-Based Learning", Paper presented at the 17^{th} EC^2 Meeting, "The Econometrics of Monetary Policy and Financial Decision-Making", December 15-16, 2006, Rotterdam. Available at `http://www.rimini.unibo.it/fanelli/fanelli\_nkpc\_learning.pdf`.

FANELLI, L. (2007b), "Testing the New Keynesian Phillips Curve Through Vector Autoregressive Models: Results from the Euro Area", *Oxford Bulletin of Economics and Statistics (forthcoming)*, available at `http://www.rimini.unibo.it/fanelli/fanelli\_nkpc.pdf`.

GALÍ, J. AND M. GERTLER (1999), "Inflation Dynamics: A Structural Econometric Analysis", *Journal of Monetary Economics*, 44, pp. 195–222.

Galí, J., G. Mm and J. D. Lopez-Salido (2001), "European Inflation Dynamics", *European Economic Review*, 45, pp. 1237–1270.

Hansen, L. P. and T. J. Sargent (1980), "Formulating and Estimating Dynamic Linear Rational Expectations Models", *Journal of Economic Dynamics and Control*, 2, pp. 7–46.

Inoue, A. and B. Rossi (2005), "Recursive Predictability Tests for Real-Time Data", *Journal of Business and Economic Statistics*, 23, pp. 336–345.

Johansen, S. (1996), *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models*, Oxford University Press, 2^{nd} revised version, Oxford.

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

Johansen, S. and A. Swensen (1999), "Testing Exact Rational Expectations in Cointegrated Vector Autoregressive Models", *Journal of Econometrics*, 93, pp. 73–91.

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, pp. 421–447.

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

Paruolo, P. (2003), "Common dynamics in I(1) systems", *Working Paper, Università dell'Insubria, Varese*, 2003/33.

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

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, pp. 133–143.

Sargent, T. J. (1999), *The Conquest of American Inflation*, Princeton University Press, Princeton.

Toda, H. Y. and T. Yamamoto (1995), "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes", *Journal of Econometrics*, 66, pp. 225–250.

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