Luca Fanelli

# Estimation of Quasi-Rational DSGE Monetary Models

Quaderni di Dipartimento

Serie Ricerche 2009, n. 3
ISSN 1973-9346

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche "Paolo Fortunati"

# Estimation of Quasi-Rational DSGE Monetary Models*

Luca Fanelli†

September 2009

**Abstract**

This paper proposes the estimation of small-scale dynamic stochastic general equilibrium (DSGE) monetary models under the quasi-rational expectations (QRE) hypothesis. The QRE-DSGE model is based on the idea that the determinate reduced form solution associated with the structural model, if it exists, must have the same lag structure as the 'best fitting' vector autoregressive (VAR) model for the observed time series. After discussing solution properties and the local identifiability of the model, a likelihood-based iterative algorithm for estimating the structural parameters and testing the data adequacy of the system is proposed. A Monte Carlo experiment shows that, even controlling for the omitted dynamics bias, the over-rejection of the nonlinear cross-equation restrictions when asymptotic critical values are used and variables are highly persistent is a relevant issue in finite samples. An application based on euro area data illustrates the advantages of using error-correcting formulations of the QRE-DSGE model when the inflation rate and the short-term interest rate are approximated as difference stationary processes. A parametric bootstrap version of the likelihood-ratio test for the implied cross-equation restrictions does not reject the estimated QRE-DSGE model.

**Keywords**: Dynamic stochastic general equilibrium model, Maximum Likelihood estimation, Quasi-Rational Expectations, VAR.

**J.E.L. Classification**: C22; C51; C52; E32; E52.

*Paper presented at the '3rd Italian Congress of Econometrics and Empirical Economics', Ancona, January 2009. A previous version of this paper circulated with the title 'Estimation of a DSGE Model Under VAR Expectations' and has been presented at the '63rd Meeting of the Econometric Society', Milano, August 2008. I wish to thank Marco Del Negro, Riccardo 'Jack' Lucchetti, Massimo Franchi and Davide Delle Monache for helpful comments on previous drafts. All errors are of my own.

†Department of Statistical Sciences, University of Bologna, via Belle Arti 41, I-40126 Bologna. e-mail: luca.fanelli@unibo.it.

# 1 Introduction

Small-scale dynamic stochastic general equilibrium (DSGE) models developed within the New Keynesian tradition, are currently treated as the benchmark of much of the monetary policy literature, given their ability to explain the impact of monetary policy on output and inflation. Despite possessing attractive theoretical properties, such as the capability of featuring potential structural sources of endogenous persistence that can account for the inertia in the data (external habit persistence, implicit indexation, adjustment costs of investment, see Christiano, Eichembaum and Evans, 2005; Smets and Wouters, 2003), DSGE models are typically rejected when compared with vector autoregressions (VAR) and have difficulties in generating sufficient endogenous persistence to match the persistence observed in the data. The empirical reliability of this class of models is an open question and misspecification remains an issue (An and Schorfheide, 2007).

From the econometric point of view, DSGE models are interpreted as inherently misspecified systems and are usually treated as restricted but parametrically incomplete representations of the actual data. Indeed, the restrictions these models impose on VARs approximating the data can be classified into two categories: (i) highly nonlinear cross-equation restrictions (CER) which involve the VAR coefficients and the structural parameters, which can be used to recover estimates of the latter; (ii) zero constraints which limit the VAR lag order. In many circumstances, the constraints in (ii) are at odds with the dynamic features and persistence observed in quarterly (monthly) time series, inducing an omitted dynamics issue that compromises the estimates of the structural parameters derived from the restrictions in (i).

Given these caveats, structural estimation and evaluation of DSGE models are feasible with standard statistical tools (maximum likelihood or Bayesian estimation)[1] once the probabilistic structure of the data has been completed with nuisance features, for instance adding dynamics (Diebold *et al.* 1998; Kozicki and Tinsley, 1999; Rudebusch, 2002*a*; 2002*b*; Fuhrer and Rudebusch 2004; Lindé, 2005, Section 5; Jondeau and Le Bihan, 2008), or manipulating arbitrarily the shock structure of the model (Smets and Wouters, 2003, 2007), or using prior distributions for the parameters with the possibility of relaxing the CER (Del Negro and Schorfheide, 2004; Del Negro et al. 2007; Del Negro and Schorfheide, 2007).[2]

[1] Aside from 'limited-information' techniques, the recent estimation of small-scale DSGE monetary models through 'full information' maximum likelihood methods include, Lindé (2005) and Cho and Moreno (2006).

[2] An alternative route has been recently explored by Cho and Moreno (2006), who focus on the small sample properties of the tests commonly used to validate the cross-equation restrictions. These authors show that the use of asymptotic critical values in samples of the sizes typically available to macroeconomists may imply false rejections of small-scale DSGE macro models.

In this paper, we propose the estimation of small-scale DSGE models under a version of the quasi-rational expectations (QRE) hypothesis, see Nerlove *et al.* (1979), Nelson and Blessler (1992), Nerlove and Fornari (1999). The extreme form of QRE simply replaces the endogenous variables appearing in the structural equations of a linear rational expectations model with their values calculated from the 'best fitting' reduced form model for them. The idea is that in a world in which the data generating process is unknown and characterized by heterogeneous information sets, rational expectations is impossible to observe and multivariate time series models as VAR models can be regarded as 'boundedly rational' predictors 'in the spirit' of rational expectations, see Branch (2004). As we show in the paper, the advantage of replacing rational expectations with QRE is that the zero restrictions in (ii) are automatically relaxed and the risk of incurring in an omitted dynamics issue is under control.

We define the QRE-DSGE model as a linear rational expectations model derived from the baseline DSGE specification whose implied reduced form solution has the same lag structure as the finite-order VAR which fits the data optimally. Once a VAR with a finite number of lags resulting from proper specification analysis is found to approximate the observed time series reasonably, the QRE-DSGE model is obtained from the baseline specification of the DSGE model by including a number of auxiliary lags of the endogenous variables, so that the associated (determinate) reduced form solution, if it exists, corresponds to a restricted version of that VAR. In this setup, the additional auxiliary parameters entering the QRE-DSGE model are not interpreted as devices required to expand dynamically a baseline theoretical specification in recognition of the observed length of real world contracts, adjustment costs, time-to-build lags, delays in information flows and decision lags, as in Kozicki and Tinsley (1999), Rudebusch (2002*a*), Fuhrer and Rudebusch (2004). Rather, the auxiliary parameters result from the assumed expectations generating system, namely from the discrepancy between the time series representation of the model under rational expectations and the time series representation of the data resulting from the agents' forecast model. The number of lags of the QRE-DSGE model is not arbitrary but is the determined from the VAR lag order.

After discussing the conditions that ensure the generic local identifiability of the structural parameters, we put forth a procedure for maximizing the likelihood function of the QRE-DSGE model, which exploits the nonlinear link between the reduced form and structural parameters iteratively.

The properties of the proposed estimation algorithm are investigated through some Monte Carlo experiments. This experiment shows that likelihood-ratio tests for the CER based on asymptotic critical values tend to over-reject the cross-equation restrictions in finite samples, providing a further explanation of the empirical failure of DSGE monetary models.

The time series upon which DSGE models are estimated are typically constructed as (or are thought of as being) deviations from steady state values. In the case of variables such as output, these are mostly log deviations from a steady state path while, for variables such as interest rates and inflation, they are level deviations from a constant steady state rate. As is known, removing a constant does not ensure stationarity if the persistence of the time series is governed by a unit root, see Cogley (2001), Fukac and Pagan (2006), Juselius and Franchi (2007), Gorodnichenko and Ng (2008) and Dees *et al.* (2008). Moreover, treating mistakenly nonstationary as stationary processes may flaw standard inferential procedures, see Johansen (2006), Li (2007), Fanelli (2008) and Fanelli and Palomba (2009). We show how the QRE-DSGE model can be transformed to account for the cointegration restrictions characterizing the observed time series.

An illustration based on euro area data for the period 1980:3-2006:4 illustrates the advantages of using error-correcting formulations of the QRE-DSGE model when the inflation rate and the short-term interest rate are highly persistent. A parametric bootstrap version of the likelihood-ratio test for the implied cross-equation restrictions does not reject the estimated QRE-DSGE model.

The rest of the paper is organized as follows. Section 2 introduces a standard formulation of small-scale DSGE monetary models and Section 3 discusses the omitted dynamics issue implicit in this class of models. Section 4 introduces formally the QRE-DSGE model and discusses identifiability of the structural parameters. Section 5 deals with the estimation algorithm and Section 6 extends the analysis to the case of cointegrated variables. The finite sample properties of the proposed estimation algorithm and the empirical size of likelihood-ration tests for the CER are studied in Section 7 on simulated data. Section 8 provides an empirical illustration based on euro area data. Some concluding remarks are provided in Section 9. Proofs and technical details are reported in the Appendix.

## 2 Model

Let $Z_t := (Z_{1,t}, Z_{2,t}, \cdots, Z_{n,t})'$ be a $n \times 1$ vector of endogenous variables and assume that the New Keynesian macroeconomic system of equations can be expressed in the form

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + c + v_t \tag{1}$$

where, $\Gamma_i := \Gamma_i(\gamma^s)$, $i = 0, f, b$ are $n \times n$ matrices of structural parameters, $c := c(\gamma^s)$ is a $n \times 1$ constant, $v_t$ is a $n \times 1$ vector which is assumed to be adapted to the sigma-field $\mathcal{F}_t$, where $\mathcal{F}_t$ represents the agents' information set at time $t$ and $E_t Z_{t+1} \equiv E(Z_{t+1} \mid \mathcal{F}_t)$. The (unrestricted) structural parameters have been collected in the $m_s \times 1$ vector $\gamma^s$.

$\Gamma_0$ is non-singular. Note that with $\Gamma_f := 0_{n\times n}$, the RE-DSGE model (1)-(3) collapses to a traditional simultaneous system of equations, whereas with $\Gamma_b := 0_{n\times n}$ one gets a 'purely forward-looking' specification.

Let $X_t := (X_{1,t}, X_{2,t}, \cdots, X_{p,t})'$ be the $p \times 1$ vector of observable variables, $p \leq n$. The measurement system

$$X_t := GZ_t \tag{2}$$

links the observable to the endogenous variables, where $G$ is $p \times n$. Throughout the paper the following assumption will be considered.

**Assumption 1** In the state-space representation (1)-(2), $G := I_p$ and $X_t$ is a 'detrended' process in the sense that the variables in $X_t$ do not embody deterministic linear trends.

Assumption 1 entails that the analysis is confined to systems in which $Z_t =: X_t$ involves observable time series and the number of shocks is equal to the number of endogenous variables ($p = n$). More general specifications can be considered without changing the main idea upon this paper is based. Assumption 1 does not rule out unit roots in $X_t$, see below.

When a direct link between the process generating $v_t$ and a set of observable 'forcing variables' is not provided by the theory, a typical completion of system (1) is obtained through the autoregressive specification

$$v_t = \Theta v_{t-1} + u_t \tag{3}$$

where $\Theta$ is a $p\times p$ (possibly diagonal) stable matrix (i.e. with eigenvalues inside the unit disk) and $u_t$ is a white noise with covariance matrix $\Sigma_u$. The specification for $v_t$ in (3) provides explicit recognition that the RE-DSGE model is not designed to capture the full extent of variation observed in the data.[3] The assumption that structural shocks are autocorrelated is common in the literature but is not generally derived from first-principles.

Under Assumption 1 the multivariate linear(ized) rational expectations model (1)-(3) nests, under precise conditions, a large class of small-scale New Keynesian models typically used in monetary policy analysis, as suggested by the example below.

**Example 1 [New Keynesian three-equation system]** Consider the following model, con-

[3] Ireland (2004, p. 1210) notes that the disturbance term $v_t$ in (1) can be interpreted as a quantity that other than soaking up measurement errors, captures all of the '*movements and co-movements in the data that the real business cycle model, because of its elegance and simplicity, cannot explain.*'

sisting in the three stylized equations:

$$y_t = \varpi_f E_t y_{t+1} + (1-\varpi_f) y_{t-1} - \delta(i_t - E_t \pi_{t+1}) + v_{1t} \quad (4)$$

$$\pi_t = \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \varrho y_t + v_{2t} \quad (5)$$

$$i_t = \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_t + \lambda_y y_t) + c_3 + v_{3t} \quad (6)$$

where $y_t$ is a measure of the output gap, $\pi_t$ the inflation rate, $i_t$ the nominal interest rate, $c_3$ a constant which is a suitable function of the desired nominal interest rate and the long run equilibrium level of inflation, and $v_{jt}$, $j = 1, 2, 3$ stochastic disturbances which can be interpreted as demand, supply and monetary shocks, respectively. The first equation, (4), is a log-linearized Euler aggregate demand (IS) curve, the second equation, (5), is the New Keynesian Phillips (NKPC) curve, and the third equation, (6), is a backward-looking Taylor type policy rule. The interested reader is referred to e.g. Clarida et al. (1999), Smets and Wouters (2003) and Christiano *et al.* (2005) for a detailed derivation of the equations in (4)-(6) and for the structural interpretation of the parameters in $\gamma^s := (\varpi_f, \delta, \gamma_f, \gamma_b, \varrho, \lambda_r, \lambda_\pi, \lambda_y, c_3)'$.[4]Referring to the notation used in (1), the model (4)-(6) is obtained, provided a measure of the output gap is available, by setting $X_t \equiv Z_t := (y_t, \pi_t, i_t)'$ and

$$\Gamma_0 := \begin{bmatrix} 1 & 0 & \delta \\ -\varrho & 1 & 0 \\ -(1-\lambda_r)\lambda_y & -(1-\lambda_r)\lambda_\pi & 1 \end{bmatrix}, \quad \Gamma_f := \begin{bmatrix} \varpi_f & \delta & 0 \\ 0 & \gamma_f & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (7)$$

$$\Gamma_b := \begin{bmatrix} (1-\varpi_f) & 0 & 0 \\ 0 & \gamma_b & 0 \\ 0 & 0 & \lambda_r \end{bmatrix}, \quad c := \begin{bmatrix} 0 \\ 0 \\ c_3 \end{bmatrix}. \blacksquare \quad (8)$$

Under Assumption 1, for some parameter configurations, the unique solution of the system (1)-(3), if it exists, can be cast in the form

$$X_t = \tilde{\Phi}_1 X_{t-1} + \tilde{\mu} + \tilde{\Psi} v_t \quad (9)$$

where the $p \times p$ matrices $\tilde{\Phi}_1 := \Phi_1(\gamma^s)$ and $\tilde{\Psi} := \Psi(\gamma^s)$ and the $p \times 1$ constant $\tilde{\mu} := \mu(\gamma^s)$ are highly nonlinear function of $\gamma^s$ and fulfill the restrictions

$$\Gamma_f (\tilde{\Phi}_1)^2 - \Gamma_0 \tilde{\Phi}_1 + \Gamma_b = 0_{p \times p}, \quad (10)$$

[4] Following Lippi and Neri (2007) one might further augment the system (4)-(5) by a money demand equation and derive the (possibly forward-looking) policy rule from the minimization of an intertemporal loss fuction. See Section 8.

$$\tilde{\mu} = (\Gamma_0 - \Gamma_f \tilde{\Phi}_1 - \Gamma_f)^{-1} c \quad (11)$$

$$vec(\tilde{\Psi}) := \left\{ [I_p \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)] - [\Theta' \otimes \Gamma_f] \right\}^{-1} vec(I_p) \quad (12)$$

where $vec(\cdot)$ is the column stacking operator and '$\otimes$' is the Kronecker product, see Binder and Pesaran (1995) and Uligh (1999).

The $\tilde{\Phi}_1$ matrix solving (10) must be real and stable for the solution (9) to be stable (asymptotically stationary) other than unique. The $\tilde{\Psi}$ matrix is non-singular.[5] The following assumption, which guarantee a unique stable solution, is considered.

**Assumption 2** In the DSGE model (1)-(3) with reduced form (9)-(12), the matrices $\Gamma_0$ and $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)$ are non-singular and the matrices $\tilde{\Phi}_1$ and $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f$ are stable.[6]

# 3 Omitted dynamics

Once the CER (10)-(12) are deduced, the structural parameters $\gamma^s$ can be estimated by maximizing one of various approximations of the likelihood function of the system. For instance, one can maximize the likelihood function of the VAR (9) subject to the restrictions (10) and (11), see Cho and Moreno (2006). Alternatively, minimum distance (MD) methods can be used to minimize the distance between the unrestricted estimates of the reduced form parameters and the structural parameters in (10)-(11), see Section 4. These procedure, however, can fail to deliver consistent estimates because of the misspecification of the VAR (9) with respect to the data. Actually, DSGE models can be misspecified in many respects, see An and Schorfheide (2007). With the term 'misspecification' here we mean the situation in which the 'correct' time series representation of $X_t$ involves more lags than the (determinate) reduced form solution associated with the RE-DSGE model. A formalized qualification is provided in Definition 1 below.

Aside from the Bayesian solution suggested by Del Negro et al. (2007), classical approaches to cope with the poor dynamic structure implied by (1)-(3) include the introduction of 'additional' dynamics in (1) to account for real-word recognition, processing, adjustment costs and time-to-build lags as in Kozicki and Tinsley (1999), Rudebusch (2002*a*, 2002*b*) and Fuhrer and Rudebusch (2004), or the manipulation of the shock structure $v_t$ as in e.g. Smets and Wouters (2003, 2007); see also Diebold *et al.* (1998).

To see how the time series structure of $v_t$ is related to the dynamic structure of the RE-DSGE model, notice that when in (3) $\Theta \neq 0_{p\times p}$, the reduced form (9) can be written as a

[5] The solution is not unique (i.e. there are multiple stable solutions) if $\tilde{\Phi}_1$ has eigenvalues inside the unit disk but the matrix $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f$ has eigenvalues outside the unit disk, see Binder and Pesaran (1995), Section 2.3.

[6] Assumption 2 ensures that the matrix $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1 - \Gamma_f)$ in (11) is non-singular.

(stable) constrained VAR of order two. Indeed, by substituting (3) into (9), using some algebra and the non singularity of $\tilde{\Psi}$, yields the expression

$$X_t = (\tilde{\Phi}_1 + \tilde{\Psi}\Theta\tilde{\Psi}^{-1})X_{t-1} - \tilde{\Psi}\Theta\tilde{\Psi}^{-1}\tilde{\Phi}_1 X_{t-2} + \tilde{\mu} + \tilde{\Psi}u_t \tag{13}$$

where $\tilde{\Phi}_1$, $\tilde{\Psi}$ and $\tilde{\mu}$ are constrained as in (10)-(12). Similarly, if $v_t$ is arbitrarily specified as a VAR of order two, the implied reduced form equilibrium of the RE-DSGE reads as a constrained VAR of order three, and so forth. Likewise, if the disturbance $v_t$ is given a mixed vector autoregressive moving average structure (VARMA), the implied reduced form can be represented as VARMA model as well, see Broze *et al.* (1990).

Consider an econometrician who observes $X_1, X_2, ..., X_T$ and finds that the 'best fitting' forecast model for $X_t$ is given by the VAR process

$$X_t = \Phi_1 X_{t-1} + \cdots + \Phi_k X_{t-k} + \mu + \varepsilon_t \tag{14}$$

where $k > 2$, $\Phi_i$, $i = 1, ..., k$ are $p \times p$ matrices of coefficients, $\mu$ a $p \times 1$ constant, and $\varepsilon_t$ is a white noise process with covariance matrix $\Sigma_\varepsilon < \infty$. Setting $\varepsilon_t := Fu_t$, with $F$ a $p \times p$ matrix, it can be easily recognized that the equations that match the agents' VAR coefficients and the reduced form of the RE-DSGE model (1)-(3) are given by

$$\Phi_1 := (\tilde{\Phi}_1 + \tilde{\Psi}\Theta\tilde{\Psi}^{-1}) \tag{15}$$

$$\Phi_2 := -\tilde{\Psi}\Theta\tilde{\Psi}^{-1}\tilde{\Phi}_1 \tag{16}$$

$$\Phi_j := 0_{p\times p} \quad , \quad j = 3, 4, ..., k \tag{17}$$

$$\mu := \tilde{\mu} \tag{18}$$

$$F := \tilde{\Psi} \tag{19}$$

where $\tilde{\Phi}_1$, $\tilde{\Psi}$ and $\tilde{\mu}$ are defined as in (10) and (12). There are two types of restrictions involved in (15)-(19): (i) the constraints in (15) and (16) and (18-(19) which define the mapping between the reduced form coefficients (the $\Pi_i$s, $\mu_\Pi$ and $F$) and the structural parameters $\gamma^s$; (ii) the zero restrictions in (17) (and in (16) when $\Theta = 0_{p\times p}$) which reduce the VAR lag order from 4 to 2 (to 1 when $\Theta = 0_{p\times p}$). It turns out that in general there exists a discrepancy between the zero restrictions in (ii) and the idea that the VAR of lag order 4 fits the data optimally. Thus the agents' best fitting model can not be regarded as the reduced form solution of the RE-DSGE model (1)-(3), unless the time series structure of $v_t$ is properly adapted. This leads to the definition below.

**Definition 1 [Omitted dynamics]** The RE-DSGE model (1)-(3) entails an omitted dynamics issue whenever the CER reduce the VAR lag order.

In the next section we define the QRE-DSGE model as a dynamic counterpart of the RE-DSGE specification (1)-(3) which circumvent the omitted dynamics issue of Definition 1.

# 4 The QRE-DSGE model

Consider the VAR for $X_t$

$$X_t = \Phi_1 X_{t-1} + \cdots + \Phi_k X_{t-k} + \mu + \varepsilon_t \ , \ \varepsilon_t \sim WN(0,\Sigma_\varepsilon) \ , \ t = 1, ..., T \tag{20}$$

where $\Phi_j$, $j = 1, ..., k$ are $p \times p$ matrices of parameters, $\mu$ is a $p \times 1$ vector of constants, $\varepsilon_t$ is a $p \times 1$ white noise process with $p \times p$ covariance matrix $\Sigma_\varepsilon$; $X_0$, $X_{-1}$, ..., $X_{-1+k}$ are fixed.

We consider the following assumptions:

**Assumption 3** The roots, $s$, of $\det[\Phi(s)] = 0$ are such that $\mid s \mid > 1$, where $\Phi(L) = I_p - \sum_{j=1}^{k} \Phi_j L^j$ is the characteristic polynomial, and $L$ is lag operator.

**Assumption 4** System (20) is the agents' forecast model with $\Phi_k \neq 0_{p\times p}$ and the coefficients $(\Phi_1, \cdots, \Phi_k, \Sigma_\varepsilon)$ are time-invariant.

Assumption 3, which is consistent with Assumption 2 when in (20) $k := 1$, rules out explosive and unit roots. We extend the analysis to the case of unit roots in Section 6. Assumption 3, in conjunction with Assumption 1, implies that we have already solved the step of linearizing the variables in $X_t$ around their deterministic/stochastic steady states.[7]

Assumption 4 implies that conditional forecasts at time $t$ are taken with respect to the sigma-field $\mathcal{H}_t := \sigma(X_1, ..., X_t) \subseteq \mathcal{F}_t$; moreover, any model restriction which reduces the VAR lag order leads to the omitted dynamics issue of the type of Definition 1. The assumption of time invariant parameters guarantees that the CER we derive below give rise to a continuos function with respect to the reduced form parameters. This assumption can be opportunely relaxed, provided the analysis we present below is applied to 'stable' sample periods.[8]

The simple adaptation of the concept of QRE (Nerlove *et al.*, 1979; Nerlove and Fornari, 1999) to the estimation of the DSGE model requires a two-step approach. In the first step, the unrestricted VAR coefficients are estimated consistently. In the second step, the implied

[7] See Dee *et al.* (2008) for a solution in which theoretically consistent measures of the steady states are obtained without resorting to 'external' procedures; see also Fukač and Pagan (2006).

[8] Many authors have shown evidence in DSGE models of the US economy of parameter instability across sample periods, especially in correspondence of changes in monetary policy regimes (Boivin and Giannoni, 2006). Misspecification tests for structural instability play a crucial role in applied research. See also Juselius and Franchi (2007).

one-step ahead forecasts

$$\widehat{E}_t X_{t+1} := \hat{\Phi}_1 X_t + \cdots + \hat{\Phi}_k X_{t-k+1} + \hat{\mu}$$

are used to replace expectations in (1), obtaining, after rearranging terms:[9]

$$(\Gamma_0 - \Gamma_f \hat{\Phi}_1) X_t = (\Gamma_f \hat{\Phi}_2 + \Gamma_b) X_{t-1}$$
$$+ \Gamma_f \hat{\Phi}_3 X_{t-2} + \cdots + \Gamma_f \hat{\Phi}_k X_{t-k+1} + (\Gamma_f \hat{\mu} + c) + v_t. \qquad (21)$$

Model (21) reads as an highly constrained simultaneous system of equations, whose structural parameters can be estimated, jointly with the equations governing the law of motion of $v_t$, with any 'limited' or 'full'-information method available in the literature.[10]

This solution has the merit of circumventing the zero restrictions on the VAR coefficients $\Phi_j$, $j = 1, ..., k$ but does not exploit the CER efficiently. To see this, note that the estimation of $\gamma^s$ in (21) by, say, full-information maximum likelihood, is asymptotically equivalent to the estimation of $\gamma^s$ obtained by a classical two-step MD method (Newey and Mcfadden, 1994) based on the minimization of the distances:

$$\begin{aligned}
(\Gamma_0 - \Gamma_f \hat{\Phi}_1)\hat{\Phi}_1 - \Gamma_f \hat{\Phi}_2 - \Gamma_b &\approx 0_{p \times p} \quad a.s. \\
(\Gamma_0 - \Gamma_f \hat{\Phi}_1)\hat{\Phi}_2 - \Gamma_f \hat{\Phi}_3 &\approx 0_{p \times p} \quad a.s. \\
&\vdots \\
(\Gamma_0 - \Gamma_f \hat{\Phi}_1)\hat{\Phi}_{k-1} - \Gamma_f \hat{\Phi}_k &\approx 0_{p \times p} \quad a.s. \\
(\Gamma_0 - \Gamma_f \hat{\Phi}_1)\hat{\Phi}_k &\approx 0_{p \times p} \quad a.s. \\
(\Gamma_0 - \Gamma_f \hat{\Phi}_1 - \Gamma_f)\hat{\mu} - c &\approx 0_{p \times 1} \quad a.s.
\end{aligned}$$

These distances have been obtained by comparing the reduced form associated with system (21) with the VAR coefficients in (20) and replacing $\Phi_j$, $j = 1, ..., k$ and $\mu$ with the unrestricted consistent estimators $\hat{\Phi}_j$ and $\hat{\mu}$ obtained in the first step.

In this paper we propose the econometric analysis of small-scale DSGE models based on a different qualification of the QRE hypothesis. We look for a specification which reconciles the

[9] If the 'best fitting' model for the data is specified as a VAR with drifting parameters and the law of motion of these parameters is associated with the time evolution of the agents' belief, the QRE is equivalent to the adaptive learning hypothesis, see Milani (2007) and references therein. In this respect, the extreme form of QRE can be interpreted as a specification of the agents' expectations formation which stands in an intermediate position between rational expectations and adaptive learning.

[10] As it is known, an important question for two-step estimators is whether the estimation of the first step affects the asymptotic variance of the second, and if so, what effect does the first step have, see Pagan (1984) and Newey and McFadden (1994).

time series approximation of the data, given by the VAR (20), with the dynamic theoretical structure implied the DSGE model, without disregarding the full set of restrictions that the latter imposes on the former. We maintain that the VAR (20), other than being the time series approximation of the data, represents the unrestricted version of the (determinate) reduced form solution associated with the DSGE model; however, to avoid the occurrence of restrictions which reduce the VAR lag order and induce the omitted dynamic bias of Definition 1, we appeal to a dynamic counterpart of the RE-DSGE model, denoted QRE-DSGE model. The QRE-DSGE model is introduced formally in Definition 2.

**Definition 2 [QR-DSGE model]** Given the VAR in (20) and the RE-DSGE model (1)-(3) with Assumptions 1-2, the QRE-DSGE model is defined as the multivariate linear rational expectations model

$$\begin{cases} \Gamma_0 X_t = \Gamma_f E_t X_{t+1} + \Gamma_b X_{t-1} + \left(\sum_{j=2}^{k} \Upsilon_j X_{t-j}\right) \mathbb{I}_{\{k\geq 2\}} + c + v_t \\ v_t = \Theta v_{t-1}(1 - \mathbb{I}_{\{k\geq 2\}}) + u_t \end{cases} \tag{22}$$

in which $\mathbb{I}_{\{\cdot\}}$ is the indicator function, $u_t$ is a white noise process with covariance matrix $\Sigma_u$ and each $\Upsilon_j$, $j = 2, ..., k$ is a $p \times p$ matrix containing auxiliary parameters.

The matrices $\Upsilon_j$s in (22) are associated with 'additional' lags of $X_t$, depending on the VAR lag order. These matrices may be specified as diagonal or not but dim($vec(\Upsilon_2 : ... : \Upsilon_k)$) should be not be too large to avoid over-parameterization. The non-zero elements in $\Upsilon_j$ are not intended to capture micro-founded propagation mechanisms or real world recognition, processing, adjustment and time-to-build lags as in e.g. Kozicki and Tinsley (1999), Rudebusch (2002*a*, 2002*b*), Fuhrer and Rudebusch (2004). Rather, the $\Upsilon_j$s are the by-product of the agents' expectations generating system in the sense that they fill the mismatch between the equilibrium implied by the RE-DSGE model and the time series approximation of the data.

Note that by construction, the QRE-DSGE model circumvent the omitted dynamics issue of Definition 1. When $k := 1$, the QRE-DSGE model (22) collapses to the RE-DSGE model (1)-(3).

The proposition that follows derives the reduced form solution of the QRE-DSGE model.

**Proposition 1 [Derivation of CER]** Under Assumptions 1-4, if a unique reduced form solution for the QRE-DSGE model (22) exists, it is given by the VAR in (20) with coefficients and disturbances subject to the restrictions $\Phi_j := \tilde{\Phi}_j$, $j = 1, ..., k$, $\mu := \tilde{\mu}$, $\varepsilon_t := \tilde{\Psi} u_t$, where the matrices $\tilde{\Phi}_j$, $\tilde{\Psi}$ and $\tilde{\mu}$ are determined by

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_1 = \Gamma_f \tilde{\Phi}_2 + \Gamma_b \tag{23}$$

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_2 = \Gamma_f \tilde{\Phi}_3 + \Upsilon_2 \tag{24}$$

$$\vdots$$

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_{k-1} = \Gamma_f \tilde{\Phi}_k + \Upsilon_{k-1}$$

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_k = \Upsilon_k$$

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1 - \Gamma_f)\tilde{\mu} = c \tag{25}$$

$$(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Psi} = I_p. \tag{26}$$

The constrained VAR is stable if the restricted companion matrix

$$\tilde{A} = \begin{bmatrix} \tilde{\Phi}_1 & \tilde{\Phi}_2 & \cdots & \tilde{\Phi}_k \\ I_p & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{p\times p} & 0_{p\times p} & I_p & 0_{p\times p} \end{bmatrix} \tag{27}$$

is stable.

**Proof**: See Appendix A.

It turns out that should the QRE-DSGE be rejected through the statistical assessment of the CER (23)-(25), the rejection can not be ascribed, in this context, to the omitted dynamics bias.

Let $\Gamma := (\Gamma^s : \Upsilon)$, $\Gamma^s := (\Gamma_0 : \Gamma_f : \Gamma_b : c)$, $\Upsilon := (\Upsilon_2 : ... : \Upsilon_k)$ be the $p \times ((k+2)p+1)$ matrix summarizing all parameters of the QRE-DSGE model. These parameters are given by the $m \times 1$ vector $\gamma = (\gamma^{s\prime}, \upsilon')'$, where $\gamma^s$ collects the $m_s$ truly structural parameters entering $\Gamma^s$, and $\upsilon$ collects the $m_\upsilon$ auxiliary parameters entering $\Upsilon$; $m = m_s + m_\upsilon$. The link between $\Gamma$ and $\gamma$ is given by

$$vec(\Gamma) = Q\gamma \tag{28}$$

where the full column rank selection matrix $Q$ has dimensions $p((k+2)p+1) \times m$.

The next proposition deals with the possibility of recovering $\gamma$ from the CER. An additional technical assumption is required.

**Assumption 5** The eigenvalues $\tilde{\varsigma}_i$, $i = 1, ...p$, of the $\tilde{\Phi}_1$ matrix are such that $\tilde{\varsigma}_i \neq \vartheta_i^{-1}$, $i = 1, ...p$, where $\vartheta_i$ are the eigenvalues of $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f$.

**Proposition 2 [Explicit form of CER and identification issues]** Consider the VAR (20) with $k \geq 2$ and the CER (23)-(25) implied by the QRE-DSGE model (22) and Assumptions 1-5. Let $\gamma_0$ be the 'true' value of $\gamma$ and $\tilde{\phi}_0$ the vector of 'true' restricted VAR coefficients, where $\tilde{\phi} := vec(\Omega)$ and $\Omega = (\tilde{\Phi}_1 : \tilde{\Phi}_2 : \cdots : \tilde{\Phi}_k : \tilde{\mu})$. (a) Necessary condition for the local identifiability of $\gamma$ is that $\dim(\gamma) = m \leq \dim(\tilde{\phi}) = a := p(pk+1)$; if $m < a$ there are $a-m$ over-identifying restrictions. (b) ; then

$$\tilde{\phi}_0 = g(\gamma_0) \tag{29}$$

where $g(\cdot)$ is a continuos differentiable function defined in a neighborhood of $\gamma_0$.

**Proof**: Appendix A.

Observe that we do not use the restrictions (26) to derive the mapping (29) as the $\tilde{\Psi}$ matrix is automatically determined from the knowledge of $\Gamma_0$, $\Gamma_f$ and $\tilde{\Phi}_1$.

## 5 Estimation

As shown in Proposition 2, the VAR (20) subject to the CER (23)-(25) can be written, in a neighborhood of $\gamma_0$, as

$$X_t = \Omega(\gamma) X_{t-1}^{*} + \varepsilon_t(\gamma) \tag{30}$$

where $\Omega(\gamma) := [\tilde{\Phi}_1(\gamma) : \cdots : \tilde{\Phi}_k(\gamma) : \tilde{\mu}(\gamma)]$, $X_{t-1}^{*} := (X_{t-1}', ..., X_{t-k}', 1)'$ and $\varepsilon_t(\gamma) := \tilde{\Psi}(\gamma) u_t$. The notation used in (30) remarks the dependence of the VAR coefficients on $\gamma$. In particular, $\tilde{\Phi}_j(\gamma) \equiv \tilde{\Phi}_j$, $j = 1, ..., k$, $\tilde{\mu}(\gamma) \equiv \tilde{\mu}$ and $\tilde{\Psi}(\gamma) \equiv \tilde{\Psi}$, where the matrices $\tilde{\Phi}_j$, $\tilde{\mu}$ and $\tilde{\Psi}$ are defined by (23)-(26).

Assuming a Gaussian distribution for $\varepsilon_t$ and denoting by $\log L(\phi)$ the concentrated log-likelihood function of the unrestricted VAR, the restricted likelihood is given by

$$\log L(\tilde{\phi}) = C - \frac{T}{2} \log \left[ \det \left( \sum_{t=1}^{T} (X_t - \Omega(\gamma) X_{t-1}^{*})(X_t - \Omega(\gamma) X_{t-1}^{*})' \right) \right] \tag{31}$$

where $\tilde{\phi} = g(\gamma)$ by (29) and $C := -\frac{pT}{2}(\log(2\pi) + 1)$.

The maximization of (31) is complicated by the fact that an analytic expression for $\Omega(\gamma)$ is not directly available.

However, the nonlinear link between the reduced form and structural parameters in (23)-(26) can be exploited iteratively as detailed in the Appendix.

Denote by $\hat{\gamma}$ be the (Q)ML estimator which maximizes the log-likelihood (31). The following proposition holds.

**Proposition 3 [Asymptotic covariance matrix of $\gamma$]** Given the VAR (20), the QRE-DSGE model (22) and the CER (23)-(26), then under the conditions of propositions 1 and 2,

$$T^{1/2}(\hat{\gamma}-\gamma_0) \underset{T\to\infty}{\overset{d}{\longrightarrow}} N(0,\mathcal{V}_\gamma) \quad , \quad \mathcal{V}_\gamma := (\mathcal{J}_\gamma' \mathcal{V}_\phi^{-1} \mathcal{J}_\gamma)^{-1} \tag{32}$$

where $\mathrm{V}_\phi = [\mathcal{I}_\infty(\phi_0)]^{-1} = [\lim_{T\to\infty} \mathcal{I}_T(\phi_0)]^{-1} := (\Sigma_\varepsilon \otimes M_{xx}^{-1})$, $M_{xx} := E(X_t^* X_t^{*\prime})$, $\mathcal{I}_\infty(\phi_0)$ is the information matrix associated with the unrestricted VAR and

$$\underset{a\times m}{\mathcal{J}_\gamma} \equiv \mathcal{J}(\gamma_0) := \left.\frac{\partial g(\gamma)}{\partial \gamma'}\right|_{\gamma:=\gamma_0} = -\underset{a\times a}{\mathcal{D}_{\phi_0}^{-1}} \times \underset{a\times m}{\mathcal{D}_{\gamma_0}} \tag{33}$$

where $a := p(pk+1)$, the $\mathcal{D}_\phi$ matrix is defined in (62) and

$$\mathcal{D}_\gamma := \underset{a\times(2p^2+a)}{\mathcal{N}_\phi} \times \underset{(2p^2+a)\times m}{Q} \tag{34}$$

with $\mathcal{N}_\phi := [\mathcal{N}_1 : \mathcal{N}_2 : \mathcal{N}_3]$ and

$$\begin{aligned} \mathcal{N}_1 &:= \left[\Omega' \otimes I_p\right] && (pk+1)\times p^2 \\ \mathcal{N}_2 &:= \left\{\left[\Omega' \tilde{\Phi}_1' \otimes I_p\right] + \left[K'\Omega' \otimes I_p\right]\right\} && (pk+1)\times p^2 \\ \mathcal{N}_3 &:= I_{(pk+1)} \end{aligned}$$

where $K$ is a $(pk+1)\times(pk+1)$ selection matrix such that $\Omega K := (0_{p\times p} : \tilde{\Phi}_2 : \cdots : \tilde{\Phi}_k : \tilde{\mu})$.

**Proof**: Appendix A.

A consistent estimate of $\mathcal{V}_\gamma$ can be obtained by replacing $\mathcal{V}_\phi$ with $\mathcal{V}_\phi := (\hat{\Sigma}_\varepsilon \otimes \widehat{\Upsilon}_{xx}^{-1})$, $\widehat{\Sigma}_\varepsilon = T^{-1}\sum_{t=1}^T (X_t - \hat{\Phi}X_{t-1}^*)(X_t - \hat{\Phi}X_{t-1}^*)'$, $\widehat{\Upsilon}_{xx} = T^{-1}\sum_{t=1}^T X_t^* X_t^{*\prime}$ and $\mathcal{J}_\gamma$ with

$$\widehat{\mathcal{J}}_\gamma = -\mathcal{D}_\phi^{-1}\mathcal{D}_\gamma \tag{35}$$

in which $\mathcal{D}_\phi$ and $\mathcal{D}_\gamma$ are obtained from the corresponding population counterparts by replacing the unknown parameters with their consistent estimates.

From Proposition 3 it follows that the asymptotic information matrix of the restricted VAR is given by

$$\mathcal{I}_\infty(\gamma_0) := \mathcal{J}_\gamma' \mathcal{I}_\infty(\phi_0) \mathcal{J}_\gamma$$

so that the local identifiability of the QRE-DSGE model is directly related to the full-columns rank condition of the Jacobian matrix $\mathcal{J}_\gamma$, see Rothenberg (1971). In principle, given a consistent estimate $\widehat{\mathcal{J}}_\gamma$, the empirical evaluation of the lack of identification of the QRE-DSGE model can be assessed by testing whether the symmetric matrix $\mathcal{J}_\gamma'\mathcal{J}_\gamma$ has zero eigenvalues.[11]

[11] It is possible to use the following result

$$T^{1/2}\left(vec(\widehat{\mathcal{J}}_\gamma'\widehat{\mathcal{J}}_\gamma) - vec(\mathcal{J}_\gamma'\mathcal{J}_\gamma)\right) \underset{T\to\infty}{\overset{d}{\longrightarrow}} N(0,\mathcal{V}_{\mathcal{J}_\gamma'\mathcal{J}_\gamma}),$$

Given the unrestricted estimates of the VAR coefficients, $\hat{\phi}$, the likelihood-ratio (LR) test for the CER is given by

$$LR := 2\left[\log L(\hat{\phi}) - \log L(g(\hat{\gamma}))\right] \underset{T\to\infty}{\overset{d}{\longrightarrow}} \chi^2(a) \tag{36}$$

where $a = \dim(\phi) - \dim(\gamma)$ is defined in Proposition 2.

# 6 Non-stationary variables

In this section we extend the analysis of the QRE-DSGE model to the case in which Assumption 3 is replaced by:

**Assumption 3'** $\det[\Phi(s)] = 0$ has exactly $p - r$ roots equal to $s = 1$, where $0 < r < p$, and the remaining roots are such that $\mid s \mid > 1$.

Assumption 3' implies that $X_t$ generated by (20) is integrated of order one (I(1)). In this case, the VAR can be represented in Vector Error Correction (VEC) form

$$\Delta X_t = \alpha\beta' X_{t-1} + \Xi W_{t-1} + \mu + \varepsilon_t \ , \ \varepsilon_t \sim WN(0,\Sigma_\varepsilon) \ , \ t = 1, ..., T \tag{37}$$

where $\alpha$ and $\beta$ are $p\times r$ matrices of full rank $r$ respectively, such that $\alpha\beta' := -(I_p - \Sigma_{j=1}^{k}\Phi_j)$, $\Xi := [\Xi_1 : \Xi_2 : ... : \Xi_{k-1}]$, $\Xi_i := -\Sigma_{j=i+1}^{k}\Phi_j$, $i = 1, ..., k-1$ and $W_{t-1} := (\Delta X_{t-1}', \Delta X_{t-2}', ..., \Delta X_{t-k+1}')'$, see Johansen (1996). For $\beta := \beta^0$, where $\beta^0$ represents an identified version of the cointegration relations, the elements in $\beta_0' X_t$ represent the stationary linear combinations of the variables in $X_t$.

For instance, considering the Example 1, if the output gap $y_t$ is the only stationary variable in $X_t := (y_t, \pi_t, i_t)'$, $r = 1$ and $\beta_0 := (1, 0, 0)'$; if also the ex-post real interest rate is stationary, $r = 2$ and $\beta_0 := (\beta_{01} : \beta_{02})$ with $\beta_{01} := (1, 0, 0)'$ and $\beta_{02} := (0, -1, 1)'$.

Once the cointegration rank $r$ has been determined from the data and the hypothesis $\beta := \beta_0$ tested and not rejected, it is possible to define the $p \times 1$ (triangular) vector

$$Y_t := \begin{pmatrix} \beta_0' X_t \\ \tau' \Delta X_t \end{pmatrix} \begin{matrix} r \times 1 \\ (p - r) \times 1 \end{matrix} \tag{38}$$

where $\tau$ is a $p \times (p - r)$ selection matrix such that $\det(\tau' \beta_{0\perp}) \neq 0$ and $\beta_{0\perp}$ is the orthogonal complement of $\beta_0$ (Johansen, 1996). In the first example above, $Y_t := (y_t, \Delta\pi_t, \Delta i_t)'$ is obtained

(which holds under the assumptions of this paper), provided an analytic expression or a suitable approximation of $V_{\mathcal{J}_\gamma' \mathcal{J}_\gamma}$ is available. Similarly, one can test the rank of $J_\gamma$ using the singular value decomposition proposed by Kleibergen and Paap (2006) (see also references therein), provided an analytic expression or a suitable approximation of the asymptotic covariance matrix of $vec(\widehat{\mathcal{J}}_\gamma)$ is derived. See Iskrev (2008) for a different approach to the evaluation of identification in DSGE models.

from (38) with $\beta_0 := (1,0,0)'$, $\tau := (e_2 : e_3)$, $e_2' := (0,1,0)$, $e_3' := (0,0,1)$; in the second example, $Y_t := (y_t, i_t - \pi_t, \Delta\pi_t)'$ is obtained with $\beta_0 := (\beta_{01} : \beta_{02})$ and $\tau := e_2$, respectively.

From the VEC (38) it turns out that the vector $Y_t$ defined in (38) admits the stable VAR representation

$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + ... + B_k Y_{t-k} + \mu^y + \varepsilon_t^y \tag{39}$$

where the matrices $B_j$, $j = 1, ..., k$ and the vector $\mu^y$ are function of the elements in $\alpha, \Xi, \mu$, $B_k$ is constrained as

$$B_k := \left[ B_{1k} \ : \ 0_{p\times(p-r)} \right] \tag{40}$$

and $\varepsilon_t^y := (\beta,\tau)'\varepsilon_t$, see Mellader *et al.* (1992) and Paruolo (2003, Theorem 2).[12]

In order to regard the system (39)-(40) as the stationary reduced-form solution associated with the QRE-DSGE model, the structural equations can be reparameterized in terms of $Y_t$, i.e. such that only the error correction terms and the changes in the variables are involved, see Fukač and Pagan (2006, Section 4.1, Strategy A). To do this, we re-write the mapping (38) as

$$Y_t := P_{\beta_0,\Delta} X_t \quad , \quad P_{\beta_0,\Delta} := \begin{pmatrix} \beta_0' \\ \tau'\Delta \end{pmatrix}$$

where $\Delta = (1 - L)$ and $P_{\beta_0,\Delta}$ is a $p \times p$ non-singular matrix; then we use $X_t := P_{\beta_0,\Delta}^{-1} Y_t$ in (22), obtaining

$$\left\{ \begin{array}{r} \Gamma_0^{\beta_0} P_{\beta_0,\Delta}^{-1} Y_t = \Gamma_f^{\beta_0} P_{\beta_0,\Delta}^{-1} E_t Y_{t+1} + \Gamma_b^{\beta_0} P_{\beta_0,\Delta}^{-1} Y_{t-1} \\ + \left( \sum_{j=2}^{k} \Upsilon_j^{\beta_0} P_{\beta_0,\Delta}^{-1} Y_{t-j} \right) \mathbb{I}_{\{k\geq 2\}} + c^{\beta_0} + v_t^{\beta_0} \\ v_t^{\beta_0} = \Theta v_{t-1}^{\beta_0} (1 - \mathbb{I}_{\{k\geq 2\}}) + u_t^{\beta_0}. \end{array} \right. \tag{41}$$

In system (41) we added a superscript '$\beta_0$' to the matrices and the shocks to remark that the over-identifying restrictions characterizing $\beta_0$ have a direct impact on the structural parameters and the shocks of the transformed structural model. Indeed, by imposing these restrictions and re-arranging the equations, system (41) can be simplified in the expression

$$\left\{ \begin{array}{r} \Gamma_0^y Y_t = \Gamma_f^y E_t Y_{t+1} + \Gamma_b^y Y_{t-1} + \left( \sum_{j=2}^{k} \Upsilon_j^y Y_{t-j} \right) \mathbb{I}_{\{k\geq 2\}} + c^y + v_t^y \\ v_t^y = \Theta v_{t-1}^y (1 - \mathbb{I}_{\{k\geq 2\}}) + u_t^y \end{array} \right. \tag{42}$$

in which the superscript '$y$' remarks that other than being formulated in terms of $Y_t$, the parameters of the QRE-DSGE model (42) accounts for all cointegration restrictions upon which $Y_t$ is defined.

[12] If the constant $\mu$ in the VEC (37) is restricted to belong to the cointegration space, i.e. $\mu := \alpha\mu_0$, then $(\beta_0' : \mu_0) \begin{pmatrix} X_t \\ 1 \end{pmatrix} = \beta_0' X_t + \mu_0$ in (38), and $\mu^y$ is zero in (39).

System (42) can be regarded as the error-correcting counterpart of the QRE-DSGE model in the presence of I(1) cointegrated variables. Note that by imposing the cointegration restrictions and re-arranging terms, the (inverse of the) difference operator $\Delta$ appearing in (41) automatically cancels out in (42). The vector of structural parameters of the 'transformed' system, $\gamma^y$, has dimension $\dim(\gamma^y) := m - f$, where $m := \dim(\gamma)$ and $f$ is number of over-identifying restrictions on $\beta_0$. Likewise, $v_t^y$ and $u_t^y$ are transformations of the original disturbances $v_t$ and $u_t$.

The example below shows how a specification of the form (42) can be obtained in practice.

**Example 2 [New Keynesian three-equation system with I(1) variables]** Turning on the Example 1, assume that $X_t$ is I(1) and is driven by a single common stochastic trend so that $r = 2$ in (37). For simplicity we set $k := 1$ in the VAR for $X_t$, hence for a given specification of $v_t$, the QRE-DSGE amounts to the RE-DSGE model (4)-(6). Assume further that $\beta := \beta_0 = (\beta_{01} : \beta_{02})$, where $\beta_{01} := (1, 0, 0)'$ and $\beta_{02} := (0, -1, 1)'$, namely the output gap and the ex-post real interest rate are stationary. This implies that there are $f = 2$ over-identifying restrictions on $\beta_0$ ($f := h - r^2$, where $h := 6$ is the total number of constraints, including normalization, on $\beta_0$). From (38) and $\tau := (0, 1, 0)'$ it turns out that $Y_t := (y_t, i_t - \pi_t, \Delta\pi_t)'$ so that

$$P_{\beta_0,\Delta} := \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & \Delta & 0 \end{bmatrix}.$$

By simple algebra, the equations (4), (5) and (6) can be reparameterized in the form

$$y_t = \varpi_f E_t y_{t+1} + (1 - \varpi_f) y_{t-1} - \delta(i_t - \pi_t) + \delta E_t \Delta\pi_{t+1} + v_{1t} \tag{43}$$

$$(i_t - \pi_t) = \lambda_r (i_{t-1} - \pi_{t-1}) + (1 - \lambda_r)(\lambda_\pi - 1)\pi_t - \lambda_r \Delta\pi_t + (1 - \lambda_r)\lambda_y y_t + c_3 + v_{3t} \tag{44}$$

$$(1 - \gamma_f - \gamma_b)\pi_t = \gamma_f E_t \Delta\pi_{t+1} - \gamma_b \Delta\pi_t + \varrho y_t + v_{2t} \tag{45}$$

where it can be recognized that the constraint $\lambda_\pi := 1$ in the policy rule (44) and the constraint $\gamma_f + \gamma_b := 1$ in the NKPC (45) are necessary to obtain a balanced system, i.e. involving only the I(0) variables in $Y_t := P_{\beta_0,\Delta}^{-1} X_t$. The counterpart of system (41) is in this case given by

$$\begin{bmatrix} 1 & \delta & \Delta^{-1}\delta \\ -\varrho & 0 & \Delta^{-1} \\ -\lambda_y (1 - \lambda_r) & 1 & \Delta^{-1}(1 - \lambda_\pi (1 - \lambda_r)) \end{bmatrix} \begin{pmatrix} y_t \\ i_t - \pi_t \\ \Delta\pi_t \end{pmatrix}$$

$$= \begin{bmatrix} \varpi_f & 0 & \Delta^{-1}\delta \\ 0 & 0 & \Delta^{-1}\gamma_f \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} E_t y_{t+1} \\ E_t(i_{t+1} - \pi_{t+1}) \\ E_t \Delta\pi_{t+1} \end{pmatrix} + \begin{bmatrix} 1 - \varpi_f & 0 & 0 \\ 0 & 0 & \Delta^{-1}\gamma_b \\ 0 & \lambda_r & \Delta^{-1}\lambda_r \end{bmatrix} \begin{pmatrix} y_{t-1} \\ i_{t-1} - \pi_{t-1} \\ \Delta\pi_{t-1} \end{pmatrix}$$

$$+\begin{pmatrix} 0 \\ 0 \\ c_3 \end{pmatrix} + \begin{pmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{pmatrix}$$

where the three matrices above correspond to $\Gamma_0^{\beta_0} P_{\beta_0,\Delta}^{-1}$, $\Gamma_f^{\beta_0} P_{\beta_0,\Delta}^{-1}$ and $\Gamma_b^{\beta_0} P_{\beta_0,\Delta}^{-1}$ in (41), respectively. By imposing $\lambda_\pi := 1$ and $\gamma_f + \gamma_b := 1$ and re-arranging the equations yields a system of the form (42) with

$$\Gamma_0^y := \begin{bmatrix} 1 & \delta & 0 \\ -\frac{\rho}{(1-\gamma_f)} & 0 & 1 \\ -(1-\lambda_r)\lambda_y & 1 & \lambda_r \end{bmatrix}, \quad \Gamma_f^y := \begin{bmatrix} \varpi_f & 0 & \delta \\ 0 & 0 & \frac{\gamma_f}{(1-\gamma_f)} \\ 0 & 0 & 0 \end{bmatrix} \tag{46}$$

$$\Gamma_b^y := \begin{bmatrix} 1-\varpi_f & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \lambda_r & 0 \end{bmatrix}, \; c^y := \begin{bmatrix} 0 \\ 0 \\ c_3 \end{bmatrix}, \; v_t^y := \begin{bmatrix} v_{1t} \\ \frac{1}{1-\gamma_f} v_{2t} \\ v_{3t} \end{bmatrix}. \tag{47}$$

The vector of structural parameters, $\gamma^y := (\varpi_f, \delta, \gamma_f, \varrho, \lambda_r, \lambda_y)'$, has dimension $\dim(\gamma^s) - f = 6$. The rejection of the two over-identifying restrictions on $\beta_0$ implies the rejection of the constraints $\lambda_\pi := 1$ and $\gamma_f + \gamma_b := 1$. ■

By replacing the cointegrated VAR with the stable system (39) and the QRE-DSGE model (22) with transformed counterpart (42), the propositions 1 and 2 can be still applied and the implied set of CER are given by

$$\begin{aligned} (\Gamma_0^y - \Gamma_f^y \tilde{B}_1)\tilde{B}_1 &= \Gamma_f^y \tilde{B}_2 + \Gamma_b^y \\ (\Gamma_0^y - \Gamma_f^y \tilde{B}_1)\tilde{B}_2 &= \Gamma_f^y \tilde{B}_3 + \Upsilon_2^y \\ &\vdots \\ (\Gamma_0^y - \Gamma_f^y \tilde{B}_1)\tilde{B}_{k-1} &= \Gamma_f^y \tilde{B}_k^* + \Upsilon_{k-1}^y \end{aligned} \tag{48}$$

$$\begin{aligned} (\Gamma_0^y - \Gamma_f^y \tilde{B}_1 - \Gamma_f^y)\tilde{\mu}^y &= c^y \\ (\Gamma_0^y - \Gamma_f^y \tilde{B}_1)\tilde{\Psi}^y &= I_p \end{aligned} \tag{49}$$

where $\tilde{B}_j$, $j = 1, ..., k$, $\tilde{\mu}^y$ are the restricted counterparts of the coefficients in (39)-(40) and $\varepsilon_t^y = \tilde{\Psi}^y u_t^y$.

The estimation of the QRE-DSGE with I(1) cointegrated variables model can be carried out as follows. If the over-identifying restrictions characterizing $\beta_0$ are not rejected, the corresponding (Q)ML estimate $\widehat{\beta}_0$ can be used in place of $\beta_0$ in (38) and treated as the 'true' value due to the super-consistency result (Johansen, 1996). Then the estimation algorithm described in Section 5 can be applied to the system (39)-(40) subject to the CER (48)-(49) to obtain the (Q)ML estimate of the vector of structural parameters $\gamma^y$.

## 7 Monte Carlo experiment

In this section, the estimation method introduced in Section 5 will be applied on simulated data to examine the efficacy of the procedure.

We consider a three equation system ($p = 3$); the reduced form is a VAR with $k = 3$ lags and Gaussian disturbances and the QRE-DSGE model is specified by taking the three equations of Example 1 as the reference model. The DGP is thus represented by the structural system

$$\Gamma_0 X_t = \Gamma_f E_t X_{t+1} + \Gamma_b X_{t-1} + \Upsilon_2 X_{t-2} + \Upsilon_3 X_{t-3} + c + u_t$$

in which $\Gamma_0$, $\Gamma_f, \Gamma_b$ and $c$ are specified as in (7)-(8) and $\Upsilon_2 := \upsilon_2 I_3$, and $\Upsilon_3 := \upsilon_3 I_3$, with $\upsilon_2$ and $\upsilon_3$ scalars; the covariance matrix of $u_t$ is specified as the identity matrix, $\Sigma_u := I_3$. The implied (determinate) reduced form is given by

$$X_t = \tilde{\Phi}_1 X_{t-1} + \tilde{\Phi}_2 X_{t-2} + \tilde{\Phi}_3 X_{t-3} + \tilde{\mu} + \tilde{\Psi} u_t \tag{50}$$

with $\tilde{\Phi}_i$, $i = 1, 2, 3$, $\tilde{\mu}$ and $\tilde{\Psi}$ determined by

$$\begin{aligned}
(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_1 &= \Gamma_f \tilde{\Phi}_2 + \Gamma_b \\
(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_2 &= \Gamma_f \tilde{\Phi}_3 + \upsilon_2 I_3 \\
(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_3 &= \upsilon_3 I_3 \\
(\Gamma_0 - \Gamma_f \tilde{\Phi}_1 - \Gamma_f)\tilde{\mu} &= c \\
(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Psi} &= I_3.
\end{aligned} \tag{51}$$

The restriction $\gamma_f + \gamma_b := 1$ is imposed in estimation. The values of the parameters $\gamma$ used to generate the data are reported in the upper panel of Table 1.To mimic situations that may occur in practise, the largest eigenvalue of the companion matrix of system (50) is set to 0.95, hence the chosen DGP entails a relatively high persistent restricted VAR. The absolute value of the largest eigenvalue of $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)$ is 0.45.

$M = 1000$ samples of length $T = 100, 200$ and 500, respectively, have been generated from the reduced form (50)-(51).[13] The ML estimates obtained with the iterative algorithm discussed in Section 5 and in the Appendix are presented in Table 1, along with the rejection frequency (empirical level) of the LR test (36) for the CER, computed using the 5% nominal critical value. The table also reports the number of times (percentage) in which the iterative procedure gives rise to estimates which violate the (asymptotic) stationarity condition (the estimated companion matrix has eigenvalues outside the unit disk) and the uniqueness condition (the

[13] All results in this section have been obtained through Ox 3.0. Results with different values of $T$ and $M$ are available upon request.

estimated $(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)$ matrix has eigenvalues outside the unit disk) which preserve determinacy of (50)-(51).

The results of this experiment show that the iterative estimation algorithm works well in practice. In particular, even with highly persistent variables, it provides a substantial control of the stability of the solution for different starting values of the structural parameters.[14]

It can be noticed, however, that with samples of size $T = 100$ the parameter $\lambda_y$, which measure the long run response of the Central Bank to output gap fluctuations, might mistakenly be considered insignificant. The situation improves substantially with the increase of the sample size. However, since in practice researchers rarely can disregard regime changes in samples of, say, $T = 200$ or more quarterly observations, the results in Table 1 support Cho and Moreno's (2006) small sample approach to inference in New Keynesian macro models.

More importantly, with samples of the sizes encountered in applications, the rejection frequency of the LR test for the CER appears sensitively higher than the significance level (9% with $T = 100$, 7.2% with $T = 200$ observations). Thus, even controlling for the omitted dynamics bias, the tests for the CER based on asymptotic critical values tend to over-reject the small-scale DSGE monetary model. This result contributes to explain, along with the results in Bekaert and Hodrick (2001), Cho and Moreno (2006) and Fanelli and Palomba (2009) obtained in related contexts, why we reject so often structural models involving forward-looking behaviour.

# 8 Empirical illustration

In this section we illustrate empirically the estimation of a QRE-DSGE model for the euro area. The reference DSGE monetary model is given by the three-equation system (4)-(6) in the Example 1.

We consider quarterly data taken from the last release of the Area-Wide Model data set described in Fagan *et al.* (2001). The variables in the vector $X_t := (\pi_t : y_t : i_t)'$ are constructed as follows. The inflation rate $\pi_t$ is measured by the log of the quarterly changes in the GDP deflator; the output gap is given by the difference between real GDP and potential output, where the latter is proxied by the HP filter applied to real GDP; $i_t$ is the short-term nominal interest rate.

Estimation is restricted to the sample 1980:3-2006:4 (which includes the initial values) so that the 'structural instability' characterizing European countries in the seventies is ruled out from the analysis. This period includes a sub-sample during which the euro area was not formally established (until the end of 1998). The 'best-fitting' model obtained from the specification

[14] The performance of the algorithm in preserving the uniqueness condition requires more investigation.

analysis is a VAR with $k := 4$ lags (hence a sample of $T := 102$ observations, net of initial values, is involved); the LR test which compares the VAR with one lag with the VAR with four lags is strongly rejected by the data (LR:=78.32, $p$-value:=0.000). Some residuals diagnostic tests for the estimated model are reported in the upper panel of Table 2. The VAR disturbances can be assumed Gaussian, though marginally.

The inflation rate and the short term interest rate appear highly persistent over the chosen sample, suggesting that $X_t$ might be approximated as an I(1) process. Table 2 reports the two largest estimated eigenvalues of the VAR companion matrix; the LR cointegration trace test; the estimated cointegration vectors $\beta := \beta_0 = (\beta_{01} : \beta_{02})$, where $\beta_{01}$ and $\beta_{02}$ are specified as in the Example 2, and the corresponding LR test for the overidentifying restrictions. The hypothesis of two cointegrating relations, $r := 2$, is not at odds with the data; note that $\beta_0$ implies that the output gap and the ex-post real interest rate are stationary up to a constant.[15]

Since the structure $\beta := \beta_0$ is not rejected (though marginally), from the discussion of Section 6 (see Example 2) and the results in Table 2 it follows that the restrictions $\lambda_\pi := 1$ and $\gamma_f + \gamma_b := 1$ must be imposed on the equations (4)-(6) for the analysis to be consistent with the stationary reduced form

$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + B_3 Y_{t-3} + B_4 Y_{t-4} + \varepsilon_t^y \tag{52}$$

in which $Y_t := (y_t - 0.0025, \Delta\pi_t, i_t - \pi_t - 0.027)'$ ($\tau := (0, 1, 0)'$) and $\varepsilon_t^y$ is a white noise process.

From Definition 2, as $k > 2$ the error-correcting counterpart of the QRE-DSGE model is represented by the system

$$\Gamma_0^y Y_t = \Gamma_f^y E_t Y_{t+1} + \Gamma_b^y Y_{t-1} + \Upsilon_2^y Y_{t-2} + \Upsilon_3^y Y_{t-3} + \Upsilon_4^y Y_{t-4} + u_t^y \tag{53}$$

where $\Gamma_0^y$, $\Gamma_f^y$ and $\Gamma_b^y$ are specified as in (46)-(47), $\Upsilon_i^y$, $i = 2, 3, 4$ are taken as diagonal and $u_t^y$ is a white noise process. The implied set of CER are summarized in the equations (48)-(49) and the vector of structural plus auxiliary parameters is defined as $\gamma^y = (\gamma^{s\prime}, \upsilon')'$, where $\gamma^s := (\varpi_f, \delta, \gamma_f, \varrho, \lambda_r, \lambda_y)$, $\upsilon := (dg(\Upsilon_2^y)', dg(\Upsilon_3^y)', dg(\Upsilon_4^y)')'$.

The upper panel of Table 3 reports the ML estimates of the structural parameters (to save space we have reported only the 'truly' structural, $\gamma^s$) with corresponding asymptotic standard errors, obtained through the iterative method illustrated in Section 5 and in the Appendix. The lower panel of Table 3 reports the LR test for the CER; in addition to the $p$-value obtained by the asymptotic distribution of the test statistic, we also computed a parametric bootstrap $p$-value associated with the test.

[15] The VAR constant is restricted to lie in the cointegration space ($\mu = \alpha\mu_0$, where $\mu_0$ is the intercept entering the cointegrating relations) since the variables do not show any linear trend.

The estimated parameters indicate both forward and backward components are important for output and inflation dynamics, partly confirming the results obtained by Lippi and Neri (2007) using different specifications of the policy rule and a shorter span of data.

While the low 't-statistics' associated with the estimated output interest elasticity ($\delta$) and the inflation elasticity to the output gap ($\varrho$) may reflect the finite sample issues documented in the simulation experiment of Section 7,[16] the considerably large standard error associated with the estimated long run response of the Central Bank to output gap fluctuations, $\lambda_y$, can be ascribed to the poor dynamic specification of the monetary policy reaction function embodied by system (53), see Lippi and Neri (2007) for details; as a result, the log-likelihood of the QRE-DSGE model tends to be relatively flact in the direction of $\lambda_y$ of the parameter space. It is worth emphasizing that the parametric bootstrap version of the LR test for the CER does not reject the estimated QRE-DSGE model for the euro area.

# 9 Concluding remarks

A growing literature attempts to 'take DSGE models to the data'. In this paper we investigate small-scale DSGE monetary models by relaxing the rational expectations hypothesis in favour of a formulation of the QRE hypothesis which enables us to circumvent the omitted dynamics bias implied by the RE-DSGE model. We have proposed the QRE-DSGE model as a specification which combines, by construction, the theoretical model with the time series features of the data. In the QRE-DSGE model, the mismatch between the reduced form solution obtained under rational expectations and the 'best fitting' VAR for the data is filled by a set of auxiliary parameters which reflect the distance between rational expectations and the agents' expectations.

A likelihood-based estimation algorithm which exploits the CER iteratively has been provided and a Monte Carlo investigation shows that the highly nonlinear nature of the restrictions may falsely led to reject the model in samples of the sizes encountered in practice. The QRE-DSGE model can be easily transformed in error-correcting form when some of the observed time series of the system can be approximated as I(1) cointegrated processes.

Application to euro area data shows that aside from the difficult task of properly specifying the monetary policy reaction, the estimation of an error-correcting formulation of a small-scale QRE-DSGE model provides reliable estimates of part of the structural parameters. A parametric

[16] Considering estimates obtained for the euro area, the estimated output interest elasticity $\delta$ is considerably lower with respect to Smets and Wouters (2003) and Lippi and Neri (2007). On the other hand, Smets and Wouters (2003) estimate a slope $\varrho$ of the NKPC of 0.007, while Lippi and Neri (2007) a slope of 0.0003. Observe, however, that while Smets and Wouters (2003) estimate a medium-scale DSGE model, the analysis in Lippi and Neri (2007) is focused on the assumption of imperfect information and discretionary monetary policy.

bootstrap version of the LR test for the CER does not reject the estimated QRE-DSGE model.

# A Appendix

## Proof of Proposition 1

If in the VAR (20) $k := 1$, the proof follows from Binder and Pesaran (1995). When $k \geq 2$, write the QRE-DSGE model in the form

$$\Gamma_0^* X_t^* = \Gamma_b^* X_{t-1}^* + \Gamma_f^* E_t X_{t+1}^* + c^* + u_t^* \tag{54}$$

where $X_t^* = (X_t', X_{t-1}', ..., X_{t-k+1}')'$, $c^* = (c', 0_{1\times p(k-1)})'$, $u_t^* = (u_t', 0_{1\times p(k-1)})'$ and

$$\Gamma_0^* = \begin{bmatrix} \Gamma_0 & 0_{p\times p} & \cdots & 0_{p\times p} \\ 0_{p\times p} & I_p & \cdots & 0_{p\times p} \\ \vdots & \vdots & \ddots & \\ 0_{p\times p} & 0_{p\times p} & \cdots & I_p \end{bmatrix}, \Gamma_f^* = \begin{bmatrix} \Gamma_f & 0_{p\times p} & \cdots & 0_{p\times p} \\ 0_{p\times p} & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \vdots & \ddots & \\ 0_{p\times p} & 0_{p\times p} & \cdots & 0_{p\times p} \end{bmatrix}$$

$$\Gamma_b^* = \begin{bmatrix} \Gamma_b & \Upsilon_2 & \cdots & \Upsilon_k \\ I_p & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{p\times p} & 0_{p\times p} & I_p & 0_{p\times p} \end{bmatrix}.$$

Write also the VAR (20) in first-order companion form

$$X_t^* = AX_{t-1}^* + \mu^* + \varepsilon_t^* \tag{55}$$

in which $\mu^* = (\mu', 0_{1\times p(k-1)})'$, $\varepsilon_t^* = (\varepsilon_t', 0_{1\times p(k-1)})'$ and

$$A = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_k \\ I_p & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{p\times p} & 0_{p\times p} & I_p & 0_{p\times p} \end{bmatrix}.$$

is stable by Assumption 3. From Binder and Pesaran (1995) it follows that provided $u_t^*$ is a bounded process, if a unique and stable solution of the system (54) exists, it takes the form (55) with $\varepsilon_t^* := \tilde{\Psi}^* u_t^*$, $A := \tilde{A}$, $\mu^* := \tilde{\mu}^*$, where $\tilde{A}$, $\tilde{\Psi}^*$and $\tilde{\mu}^*$ are determined by the restrictions

$$\Gamma_f^*(\tilde{A})^2 - \Gamma_0^*\tilde{A} + \Gamma_b^* = 0_{pk\times pk} \tag{56}$$

$$\tilde{\mu}^* = (\Gamma_0^* - \Gamma_f^*\tilde{A} - \Gamma_f^*)^{-1} c^* \tag{57}$$

$$\tilde{\Psi}^* = (\Gamma_0^* - \Gamma_f^*\tilde{A})^{-1}. \tag{58}$$

Uniqueness and stability obtains if $\Gamma_0^*$ and $(\Gamma_0^* - \Gamma_f^* \tilde{A})$ are non-singular and $(\Gamma_0^* - \Gamma_f^* \tilde{A})^{-1}\Gamma_f^*$ and $\tilde{A}$ are stable, respectively. Observe that the stability of the unrestricted VAR companion matrix $\Phi^*$ is not sufficient for the stability of $\tilde{A}$.

$\Gamma_0^*$ is non-singular if $\Gamma_0$ is non-singular. Moreover, the matrix

$$(\Gamma_0^* - \Gamma_f^* \tilde{A}) = \begin{bmatrix} \Gamma_0 - \Gamma_f \tilde{\Phi}_1 & -\Gamma_f \tilde{\Phi}_2 & \cdots & -\Gamma_f \tilde{\Phi}_k \\ 0_{p\times p} & I_p & \cdots & 0_{p\times p} \\ \vdots & \vdots & \ddots & \\ 0_{p\times p} & 0_{p\times p} & \cdots & I_p \end{bmatrix}$$

is non-singular since the sub-matrix in the upper-left corner is non-singular (Assumption 2). Finally, by using inversion formulas for partitioned matrix, one gets

$$(\Gamma_0^* - \Gamma_f^* \tilde{A})^{-1}\Gamma_f^* = \begin{bmatrix} (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f & 0_{p\times p} & \cdots & 0_{p\times p} \\ 0_{p\times p} & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \vdots & \ddots & \\ 0_{p\times p} & 0_{p\times p} & \cdots & 0_{p\times p} \end{bmatrix}$$

and this matrix has stable eigenvalues if the sub-matrix in the upper-left corner has stable eigenvalues, condition guaranteed by Assumption 2. It turns out that the unique solution to the multivariate linear rational expectations model (54) takes the form (55) with coefficients subject to the restrictions (56)-(58). Using the definition of the variables, the expressions in (56)-(58) are equivalent to (23)-(26). If the $\tilde{A}$ matrix solving (56) is real and stable, the solution is stable, other than unique. Assumptions 4 guarantees that $\Upsilon_k \neq 0_{p\times p}$ in (23)-(26). ■

## Proof of Proposition 2

Defined $\Omega := (\tilde{\Phi} : \widetilde{\mu})$, write the CER (23)-(25) compactly as

$$\Gamma_0 \Omega - \Gamma_f \tilde{\Phi}_1 \Omega - \Gamma_f \Omega K - \breve{\Gamma}_b = 0_{p\times(pk+1)} \tag{59}$$

where $K$ is a $(pk+1)\times(pk+1)$ selection matrix such that $\Omega K := (0_{p\times p} : \tilde{\Phi}_2 : \cdots : \tilde{\Phi}_k : \tilde{\mu})$ and $\breve{\Gamma}_b := (\Gamma_b : c : \Upsilon)$. Let $\tilde{\phi} := vec(\Omega)$ and $\tilde{\phi}_j := vec(\tilde{\Phi}_j)$, $j = 1, ..., k$; obviously, $\tilde{\phi} \equiv (\tilde{\phi}_1', ..., \tilde{\phi}_k', \tilde{\mu}')'$. By applying the vec operator to both sides of (59) we obtain the vector function

$$f(\tilde{\phi}, \gamma) = (I_a \otimes \Gamma_0)\tilde{\phi} - (I_a \otimes \Gamma_f \tilde{\Phi}_1)\tilde{\phi}$$
$$- (K' \otimes \Gamma_f)\tilde{\phi} - vec(\breve{\Gamma}_b) = 0_{a\times 1} \tag{60}$$

where $a := p(pk+1)$ and $f : \mathbb{S} \rightarrow \mathbb{R}^a$ is continuous differentiable on the set $\mathbb{S}$ of $\mathbb{R}^{a+m}$ (Assumption 4). If the $a \times a$ Jacobian matrix

$$\mathcal{D}_\phi := \frac{\partial f(\tilde{\phi}, \gamma)}{\partial \tilde{\phi}'} \tag{61}$$

is non-singular at $(\tilde{\phi}_0, \gamma_0)$, by the implicit function theorem, $\tilde{\phi}$ can be uniquely expressed as function of $\gamma$ in a neighborhood $\mathcal{N}(\gamma_0) \subset \mathbb{R}^m$ of $\gamma_0$. In particular, $\tilde{\phi} = g(\gamma)$ for all $\gamma$ in $\mathcal{N}(\gamma_0)$, where $g$ is the differentiable function $g : \mathcal{N}(\gamma_0) \rightarrow \mathbb{R}^a$. The order condition $m \leq a$ guarantees that the number of structural parameters in $\gamma$ does not exceed the number of reduced form coefficients; if $m < a$, there are $a - m$ over-identifying restrictions. This proves part (a).

To prove (b) we show that the Jacobian in (61) is non-singular at $(\tilde{\phi}_0, \gamma_0)$. To compute the matrix $\mathcal{D}_\phi$ we apply the vec operator to each set of restrictions in (23)-(26) and decompose $\mathcal{D}_\phi$ in blocks of dimension $p^2 \times p^2$. For instance, taking the vec of (23) and deriving the resulting expression with respect to the vector $\tilde{\phi}_1$ one gets

$$\begin{aligned}\mathcal{D}_\phi^{1,1} &:= \frac{\partial vec[(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_1 - \Gamma_f \tilde{\Phi}_2 - \Gamma_b]}{\partial vec(\tilde{\Phi}_1)'} \\ &= (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f)[(\tilde{\Phi}_1' \otimes I_p) + (I_p \otimes \tilde{\Phi}_1)]\end{aligned}$$

where $\mathcal{D}_\phi^{1,1}$ is the left-upper block of $\mathcal{D}_\phi$. Similarly, by deriving $vec[(\Gamma_0 - \Gamma_f \tilde{\Phi}_1)\tilde{\Phi}_1 - \Gamma_f \tilde{\Phi}_2 - \Gamma_b]$ with respect to $\tilde{\phi}_l$, $l = 2, ..., k$ and $\tilde{\mu}$ yields

$$\begin{aligned}\mathcal{D}_\phi^{1,2} &: = -(I_p \otimes \Gamma_f); \\ \mathcal{D}_\phi^{1,l} &: = 0_{p^2 \times p^2} \ , \ \ l = 2, ..., k \\ \mathcal{D}_\phi^{1,k+1} &: = 0_{p^2 \times p}.\end{aligned}$$

It is assumed that all derivatives are evaluated at the 'true' parameter values $(\widetilde{\phi}_0, \gamma_0)$. Likewise, by taking the vec of (24) and deriving the resulting expression with respect to $\tilde{\phi}_l$, $l = 1, ..., k$ and $\tilde{\mu}$ yields:

$$\begin{aligned}\mathcal{D}_\phi^{2,1} &:= -(\tilde{\Phi}_2' \otimes \Gamma_f); \\ \mathcal{D}_\phi^{2,2} &:= (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f \tilde{\Phi}_1); \\ \mathcal{D}_\phi^{2,3} &:= -(I_p \otimes \Gamma_f); \\ \mathcal{D}_\phi^{2,j} &: = 0_{p^2 \times p^2} \ \ , \ \ l = 3, ..., k \\ \mathcal{D}_\phi^{3,k+1} &:= 0_{p^2 \times p},\end{aligned}$$

and so forth. The $\mathcal{D}_\phi$ matrix reads as

$$\mathcal{D}_\phi := \begin{bmatrix} \mathcal{D}_\phi^{1,1} & -(I_p \otimes \Gamma_f) & 0_{p^2 \times p^2} & 0_{p^2 \times p^2} & \cdots & 0_{p^2 \times p} \\ -(\tilde{\Phi}_2' \otimes \Gamma_f) & \mathcal{D}_\phi^{2,2} & -(I_p \otimes \Gamma_f) & 0_{p^2 \times p^2} & \cdots & 0_{p^2 \times p} \\ -(\tilde{\Phi}_3' \otimes \Gamma_f) & 0_{p^2 \times p^2} & \mathcal{D}_\phi^{3,3} & -(I_p \otimes \Gamma_f) & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -(\tilde{\Phi}_k' \otimes \Gamma_f) & 0_{p^2 \times p^2} & 0_{p^2 \times p^2} & \cdots & \mathcal{D}_\phi^{k,k} & 0_{p^2 \times p} \\ -(\tilde{\mu}' \otimes \Gamma_f) & 0_{p \times p^2} & 0_{p \times p^2} & \cdots & 0_{p \times p^2} & \mathcal{D}_\phi^{k+1,k+1} \end{bmatrix} \quad (62)$$

where $\mathcal{D}_{\phi}^{l,l} := (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f \tilde{\Phi}_1')$, $l = 2, ..., k$, and $\mathcal{D}_{\phi}^{k+1,k+1} := (\Gamma_0 - \Gamma_f \tilde{\Phi}_1' - \Gamma_f)$. By construction, each block of $p^2$ columns in $\mathcal{D}_{\phi}$ can not be obtained as linear combinations of the other blocks of $p^2$ columns. The $\mathcal{D}_{\phi}^{l,l}$, $l = 2, ..., k$ matrices on the main diagonal are non-singular by Assumption 2. Since $\mathcal{D}_{\phi}^{k+1,k+1} = (\Gamma_0 - \Gamma_f \tilde{\Phi}_1) - \Gamma_f = (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)[I_p - (\Gamma_0 - \Gamma \tilde{\Phi}_1)^{-1}\Gamma_f]$, also this matrix is non-singular by Assumption 2. Finally,

$$\begin{aligned}
&\mathcal{D}_{\phi}^{1,1} := (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f)[(\tilde{\Phi}_1' \otimes I_p) + (I_p \otimes \tilde{\Phi}_1)] \\
= \quad & [I_p \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)] - (\tilde{\Phi}_1' \otimes \Gamma_f) \\
= \quad & [I_p \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)][I_{p^2} - (I_p \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1})(\tilde{\Phi}_1' \otimes \Gamma_f)] \\
= \quad & [I_p \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)][I_{p^2} - (\tilde{\Phi}_1' \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f)]
\end{aligned}$$

hence the first term in the product is non-singular by Assumption 2, whereas the second matrix is non-singular as the eigenvalues of the matrix $\tilde{\Phi}_1' \otimes (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1}\Gamma_f$ are different from 1 by Assumption 5. We have therefore proved that each block of $p^2$ columns in (62), other than being linearly independent from the other block, has rank $p^2$. Accordingly $rank(\mathcal{D}_{\phi}) = kp^2 + p = a$. This proves part $(b)$. ■

## Proof of Proposition 3

Let $\log L(\phi)$ be the concentrated log-likelihood function of the unrestricted VAR and $\log L(g(\gamma))$ the concentrated log-likelihood of the VAR subject to the restrictions (29). It is known, under Assumption 3, that

$$T^{1/2}(\hat{\phi} - \phi_0) \underset{T\to\infty}{\xrightarrow{d}} N(0, \mathcal{V}_{\phi}) \quad , \quad \mathcal{V}_{\phi} = [\mathcal{I}_{\infty}(\phi_0)]^{-1} := \Sigma_{\varepsilon} \otimes M_{xx}$$

implies

$$T^{1/2}(\hat{\gamma} - \gamma_0) \underset{T\to\infty}{\xrightarrow{d}} N(0, \mathcal{V}_{\gamma}) \quad , \quad \mathcal{V}_{\gamma} := (\mathcal{J}_{\gamma}' \mathcal{V}_{\phi}^{-1} \mathcal{J}_{\gamma})^{-1}$$

where $\mathcal{J}_{\gamma} := \frac{\partial g(\gamma)}{\partial \gamma'}$ is the Jacobian matrix associated with the function (29) evaluated at $\gamma_0$, whose existence id guaranteed by Proposition 2. Using (60) and the implicit function theorem, it turns out that

$$\mathcal{J}_{\gamma} := -\mathcal{D}_{\phi}^{-1} \times \mathcal{D}_{\gamma} \tag{63}$$

where $\mathcal{D}_{\phi}$ is the Jacobian matrix defined in (62) and

$$\mathcal{D}_{\gamma} := \frac{\partial f(\widetilde{\phi}, \gamma)}{\partial \gamma'} = \frac{\partial f(\widetilde{\phi}, \gamma)}{\partial vec(\Gamma)'} \times \frac{\partial vec(\Gamma)}{\partial \gamma'} = \frac{\partial f(\widetilde{\phi}, \gamma)}{\partial vec(\Gamma)'} Q. \tag{64}$$

To evaluate $\frac{\partial f(\widetilde{\phi},\gamma)}{\partial vec(\Gamma)'}$ we re-write the function

$$f(\tilde{\phi},\gamma) := vec[\Gamma_0\Omega - \Gamma_f\tilde{\Phi}_1\Omega - \Gamma_f\Omega K - \breve{\Gamma}_b]$$

in the alternative format

$$\begin{aligned} f(\tilde{\phi},\gamma) &\equiv [\Omega' \otimes I_p]vec(\Gamma_0) \\ &- \left\{[\Omega'\tilde{\Phi}_1' \otimes I_p] + [K'\Omega' \otimes I_p]\right\} vec(\Gamma_f) - vec(\breve{\Gamma}_b) \\ &= [\mathcal{N}_1 : \mathcal{N}_2 : \mathcal{N}_3] \begin{pmatrix} vec(\Gamma_0) \\ vec(\Gamma_f) \\ vec(\breve{\Gamma}_b) \end{pmatrix} = \mathcal{N}_\phi vec(\Gamma) \end{aligned}$$

where

$$\begin{aligned} \mathcal{N}_1 &:= \left[\Omega' \otimes I_p\right] & & p(pk+1) \times p^2 \\ \mathcal{N}_2 &:= \left\{\left[\Omega'\tilde{\Phi}_1' \otimes I_p\right] + \left[K' \begin{pmatrix} \tilde{\Phi}' \\ \tilde{\mu}' \end{pmatrix} \otimes I_p\right]\right\} & & p(pk+1) \times p^2 \\ \mathcal{N}_3 &:= I_a. \end{aligned}$$

It turns out that

$$\frac{\partial f(\widetilde{\phi},\gamma)}{\partial vec(\Gamma)'} \equiv \mathcal{N}_\phi.$$

Note that the number of columns of $\mathcal{N}_\phi$, $2p^2 + a$, is equal to the number of rows of $Q$, $p((k+2)p+1)$; the Jacobian matrix (64) is thus given by

$$\mathcal{D}_\gamma := \mathcal{N}_\phi \; Q$$

as in (34).■

## Estimation algorithm

Before discussing the suggested estimation algorithm, consider the constrained VAR in (30). Let $S$ be a $p \times p$ positive definite matrix. Any vector $\hat{\gamma}(S)$ which minimizes the quadratic form

$$R_T(\gamma) := \frac{1}{T}\sum_{t=1}^{T}(X_t - \Omega(\gamma)X_{t-1}^*)' \; S \; (X_t - \Omega(\gamma)X_{t-1}^*) \tag{65}$$

given the observation $X_1, X_2, ..., X_T$ is called a MD estimator of $\gamma_0$, see Phillips (1976) and Malinvaud (1980, Ch. 9). These authors shows that, under a set of regularity conditions, if one iterates the MD estimator $\hat{\gamma}(\cdot)$ by replacing $S$ at each iteration with the inverse of the moment matrix of residuals obtained in the previous iteration, the algorithm based on the iterative minimization of (65) will converge eventually to the quasi-ML (QML) estimator that maximizes (31).[17]

This result inspires the iterative likelihood-based estimation method sketched below. Observe that the CER in (23)-(26) combine two type of restrictions. (23) involves a quadratic matrix equation of the form

$$H_1 V^2 + H_2 V + H_3 = 0_{p\times p} \tag{66}$$

where the matrices $H_i, i = 1, 2, 3$ contain the structural parameters and $V$ contains the coefficients associated with the first lag of the VAR. $H_2$ is non-singular while $H_1$ and $H_3$ can be singular. The remaining set of nonlinear restrictions (24)-(26) do not involve quadratic matrix equations. However, as shown in the proof of Proposition 1, a compact representation of (23)-(26) is given by (56)-(57), where in particular (56) is of the form (66).

Assume temporarily that $H_i, i = 1, 2, 3$ are known (i.e. $\gamma$ is known). Then the solution $V$ to (66) can be derived from the 'theory of solvents' and is equivalent to a quadratic eigenvalue problem (Higham and Kim, 2000) which in turn can be re-formulated in terms of a generalized eigenvalue problem involving the matrices

$$\begin{bmatrix} 0_{p\times p} & I_p \\ -H_3 & -H_2 \end{bmatrix} \quad , \quad \begin{bmatrix} I_p & 0_{p\times p} \\ 0_{p\times p} & H_1 \end{bmatrix}$$

see Uhlig (1999). In general, equation (66) can have no solution, a finite positive number, or infinitely many. The solution can be a complex matrix. A natural way to get an approximate real matrix solution to (66) is to appeal to numerical solutions based on function iterations (Higham and Kim, 2000).

Re-write the CER in (59) as $\Omega = W(\Omega, \gamma)$, where $W(\cdot)$ is matrix continuous matrix function, see below. Then define the iterations

$$\Omega_j := W(\Omega_{j-1}, \gamma) \quad , \quad j = 1, 2, ... \tag{67}$$

where $\Omega_0$ is an initial guess for $\Omega$. These iterations provide an approximate solution which corresponds to a real matrix provided $\Omega_0$ is real. Unfortunately, the iterations they can not in general be transformed to a simpler form and so it is difficult to obtain convergence results of

[17] Alternatively, MD estimation of the QR-DSGE model can be based on the implicit form representation of the CER, $f(\phi, \gamma) = 0$, see equation (60). Given the unrestricted estimates of the VAR coefficients obtained in the first step, $\hat{\phi}$, the method is based on the criterion

$$\min_{\gamma} f(\hat{\phi}, \gamma)' S \, f(\hat{\phi}, \gamma)$$

for an oppropriate choice of $S$, see e.g. Newey and Mcfadden, (1994).

practical applicability. However, if $\gamma$ is unknown, at each iteration the left-hand side of (67) depends only on $\gamma$ through the two alternative expressions

$$\begin{aligned} \Omega_j \quad &: \quad = W(\Omega_{j-1}, \gamma) \\ &: \quad = \Gamma_0^{-1}\Gamma_f \tilde{\Phi}_1 \Omega_{j-1} + \Gamma_0^{-1}\Gamma_f \Omega_{j-1} K + \Gamma_0^{-1} \breve{\Gamma}_b; \end{aligned} \tag{68}$$

$$\begin{aligned} \Omega_j \quad &: \quad = W(\Omega_{j-1}, \gamma) \\ &= \quad \Gamma_f \Omega_{j-1} K (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1} + \breve{\Gamma}_b (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1} \end{aligned} \tag{69}$$

The suggested estimation algorithm starts from the unrestricted consistent estimates of the VAR coefficients obtained with OLS (Gaussian ML), i.e. $\Omega_0 := \hat{\Omega} := [\hat{\Phi}_1 : \hat{\Phi}_2 : \ldots : \hat{\Phi}_k : \hat{\mu}]$, and then uses at each iteration (68) or (69) to express the VAR coefficients as (unique) function of the structural parameters. At the $j$th iteration, a (Q)ML estimate of $\gamma$, $\hat{\gamma}^j$, is obtained by maximizing the log-likelihood (31) in which $\Omega(\gamma)$ is replaced with $\Omega_j$ in (68) or (69). Newton or Quasi-Newton techniques such as the BFGS method (Fletcher,1987) can be used at each iteration and $\hat{\Sigma}_\varepsilon^j = T^{-1} \sum_{t=1}^{T} (X_t - \hat{\Omega}_j X_{t-1}^*)(X_t - \hat{\Omega}_j X_{t-1}^*)'$ is the corresponding estimate of the constrained VAR disturbances covariance matrix, where $\hat{\Omega}_j$ is obtained from (68) or (69) by replacing $\gamma$ with $\hat{\gamma}^j$. The procedure is iterated until some convergence criterion is met.

We have not a formal proof of convergence. Our conjecture is that it should be proved that the algorithm converges under the same set of conditions Phillips (1976) employs to prove that the iterated MD estimator based on the minimization of (65) converges (and attains the QML estimator). The simulation experiment reported in Section 7 shows that convergence is achieved, on average, in 27-29 iterations; moreover, the procedure delivers an accurate solution to the quadratic matrix equation (23).

As it stands, the iterative algorithm does not impose neither uniqueness (i.e. the eigenvalues of the matrix $(\Gamma_0 - \Gamma_f \ \tilde{\Phi}_1)^{-1}\Gamma_f$ must lie inside the unit disk), nor stability (i.e. the eigenvalues of the restricted VAR companion matrix $\tilde{A}$ in (27) must lie inside the unit disk). Nevertheless, the iterations can be stopped as soon as violation of one of the two conditions or both, is detected. The simulation experiment shows that for data generating processes in which the absolute value of the largest eigenvalue of $\tilde{A}$ is specified close to one, the suggested algorithm tends to preserve stability, see Table 1. The experiment also shows that the procedure provides a strict control of uniqueness when the absolute value of the largest eigenvalue of $(\Gamma_0 - \Gamma_f \ \tilde{\Phi}_1)^{-1}\Gamma_f$ is relatively far from one, see Table 1.

## References


An, S., Schorfheide, F. (2007), Bayesian analysis of DSGE models, *Econometric Reviews* 26, 113-172.

Bekaert G., Hodrick R. (2001), Expectations hypotheses tests, *Journal of Finance* 56, 1357-1394.

Binder, M. and Pesaran, M. H. (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*, pp. 139-187 (Ch. 3). Oxford: Blackwell.

Boivin, J., Giannoni, M.P. (2006), Has monetary policy become more effective ?, *Review of Economics and Statistics* 88, 445-462.

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

Broze, L., Gourieroux, C. and Szafar, A. (1990), *Reduced forms of rational expectations models*, Harwood Academic Publishers, New York.

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

Cho, S., Moreno, A. (2006), A small-sample study of the New-Keynesian macro model, *Journal of Money Credit and Banking* 38, 1462-1482.

Christiano, L.J., Eichembaum, M., Evans, C.L. (2005), Nominal rigidities and the dynamic effects of a shock to monetary policy, *Journal of Political Economy* 113, 1-45.

Clarida, R., Gali, J., Gertler, M. (1999), The science of monetary policy: a New Keynesian perspective, *Journal of Economic Literature* 37, 1661-1707.

Cogley, T. (2001), estimating and testing rational expectations models when the trend specification is uncertain, J*ournal of Economic Dynamics and Control* 25, 1485-1525.

Dees, S., Pesaran, H.M., Smith, V., Smith, R.P. (2008), Identification of New Keynesian Phillips curves from a global perspective, *IZA Discussion Paper* No. 3298.

Del Negro, M., Schorfheide, F. (2004), Priors from general equilibrium models for VARs, *International Economic Review* 45, 643-673.

Del Negro, M., Schorfheide, F. (2009), Monetary policy with potentially misspecified models, *American Economic Review*, forthcoming. *NBER Working Paper* No. 13099.

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R. (2007), On the fit of New Keynesian models, *Journal of Business and Economic Statistics* 25, 123-143.

Diebold, F. X., Ohanian, L. E., Berkowitz, J. (1998), Dynamic equilibrium economies: a framework for comparing models and data, *Review of Economic Studies* 65, 433-452.

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), 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., Palomba, G. (2009), Simulation-based tests of forward-looking models under VAR learning dynamics, *Journal of Applied Econometrics*, forthcoming.

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

Fuhrer, J., Rudebusch, G.D. (2004), Estimating the Euler equation for output, *Journal of Monetary Economics* 51, 1133-1153.

Fukač, M., Pagan, A. (2006), Limited information estimation and evaluation of DSGE models, *NCER Working Paper* No. 6.

Gorodnichenko, Y., Ng, S. (2008), Estimation of DSGE models when the data are persistent, mimeo.

Higham, N.J., Kim, H-M. (2000), Numerical analysis of a quadratic matrix equation, *IMA Journal of Numerical Analysis* 20, 499-519.

Ireland, P.N. (2004), A method for taking models to the data, *Journal of Economic Dynamics and Control* 28, 1205-1226.

Iskrev, N. (2008), How much do we learn from the estimation of DSGE models ? A case study of identification issues in a New Keynesian business cycle model, mimeo.

Johansen, S. (1996). *Likelihood Based Inference in Cointegrated Vector Autoregressive Models.* 2nd edn. Oxford: Oxford University Press.

Johansen, S. (2006), Confronting the economic model with the data. In D. Colander (ed), *Post Walrasian Macroeconomics*, Cambridge University Press, Cambridge MA.

Jondeau, E., Le Bihan, H. (2008), Examining bias in estimators of linar rational expectations models under misspecification, *Journal of Econometrics* 143, 375-395.

Juselius, K., Franchi, M. (2007), Taking a DSGE model to the data meaningfully, *Economics: The Open-Access, Open-Assessment E-Journal* 1, 2007-4.

Kleibergen, F., Paap, R. (2006), Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* 133, 97-126.

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

Li, H. (2007), Small-sample inference in rational expectations models with persistent data, *Economic Letters* 95, 203-210.

Lippi, F., Neri, S. (2005), Information variables for monetary policy in an estimated structural model of the euro area, *Journal of Monetary Economics* 54, 1256-1270.

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

Malinvaud, E. (1980), *Statistical methods of econometrics*, North-Holland, Amsterdam.

Mellader, E., Vredin, A. and Warne, A. (1992), Stochastic trends and economic fluctuations in a small open economy. *Journal of Applied Econometrics 7*, 369-394.

Milani, F. (2007), Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics* 54, 2065-2082.

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

Nelson, R. G., Bessler, D. A. (1992), Quasi-rational expectations: Experimental evidence, *Journal of Forecasting* 11, 141-156.

Nerlove, M., Fornari, I. (1999), Quasi-rational expectations, an alternative to fully rational expectations: An application to US beef cattle supply, *Journal of Econometrics* 83, 129-161.

Nerlove, M., Grether, D., Carvalho, J.L. (1979), *Analysis of Economic Time Series*. Academic Press, New York.

Newey, W.K., McFadden, D. (1994), Large sample estimation and hypothesis testing. In R.F. Engle and D. McFadden (eds.), *The Handbook of Econometrics*, pp. 2111-2245 (Vol. 4). North Holland.

Pagan, A. (1984), Econometric issues in the analysis of regressions with generated regressors, I*nternational Economic Review* 25, 221-247.

Paruolo, P. (2003), Common dynamics in I(1) systems, Working Paper No. 2003/33, Università dell'Insubria, Varese.

Pesaran, H. M. (1981), Identification in rational expectation models, *Journal of Econometrics* 16, 375-398.

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

Phillips, P.C.B. (1976), The iterated minimum distance estimator and the quasi-maximum likelihood estimator, *Econometrica* 44, 449-460.

Rothenberg, T. (1971), Identification in parametric models, *Econometrica* 39, 577-591.

Rudebusch, G.D. (2002*a*), Assessing nominal income rules for moetary policy with model and data uncertainty, *Economic Journal* 112, 402-432.

Rudebusch, G.D. (2002*b*), Term structure evidence on interest rate smoothing and monetray policy inertia, *Journal of Monetary Economics* 49, 1161-1187.

Smets, F., Wouters, R. (2003), An estimated dynamic stochastic general equilibrium model of the euro area, *Journal of the European Economic Association* 1, 1123-1175.

Smets, F., Wouters, R. (2007), Shocks and frictions in U.S. business cycles, *ECB Working Paper* No. 722.

Uhlig, H. (1999), A Toolkit for analyzing nonlinear dynamic stochastic models easily. In R. Marimon and A. Scott (eds.), *Computational methods for the study of dynamics economies*, pp. 30-61, Oxford University Press, New York.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| True values in DGP: | | | | | | | | | |
| $\varpi_f$ | $\delta$ | $\gamma_f$ | $\varrho$ | $\lambda_r$ | $\lambda_\pi$ | $\lambda_y$ | $c_3$ | $\omega_2$ | $\omega_3$ |
| 0.25 | 0.10 | 0.30 | 0.13 | 0.5 | 1.5 | 0.5 | 0.4 | -0.3 | -0.1 |
| Absolute value of largest eigen. of restricted VAR companion matrix $\tilde{\Phi}^*$: 0.95 | | | | | | | | | |
| Absolute value of largest eigen. of $(\Gamma_0 - \Gamma_f\ \tilde{\Phi}_1)^{-1}\Gamma_f$: 0.45 | | | | | | | | | |
| ML estimates | | | | | | | | | |
| | | | | | $T$=100 | | | | |
| 0.252 | 0.103 | 0.306 | 0.132 | 0.486 | 1.503 | 0.501 | 0.399 | -0.291 | -0.103 |
| (0.063) | (0.057) | (0.058) | (0.066) | (0.080) | (0.312) | (0.367) | (0.103) | (0.043) | (0.024) |
| (0.067)* | (0.057)* | (0.057)* | (0.085)* | (0.075)* | (0.205)* | (0.222)* | (0.110)* | (0.068)* | (0.044)* |
| Frequency of rejections of LR test for CER: 0.09 (nominal level: 0.05) | | | | | | | | | |
| # of times in which stability is violated : 0 ; # of times in which uniqueness is violated : 0 | | | | | | | | | |
| | | | | | $T$=200 | | | | |
| 0.250 | 0.103 | 0.302 | 0.132 | 0.492 | 1.502 | 0.496 | 0.399 | -0.295 | -0.102 |
| (0.043) | (0.039) | (0.041) | (0.046) | (0.056) | (0.218) | (0.258) | (0.103) | (0.030) | (0.017) |
| (0.047)* | (0.038)* | (0.040)* | (0.058)* | (0.054)* | (0.146)* | (0.153)* | (0.110)* | (0.049)* | (0.031)* |
| Frequency of rejections of LR test for CER: 0.072 (nominal level: 0.05) | | | | | | | | | |
| # of times in which stability is violated : 0 ; # of times in which uniqueness is violated : 0 | | | | | | | | | |
| | | | | | $T$=500 | | | | |
| 0.248 | 0.101 | 0.301 | 0.131 | 0.497 | 1.502 | 0.497 | 0.397 | -0.298 | -0.102 |
| (0.027) | (0.024) | (0.025) | (0.029) | (0.035) | (0.136) | (0.163) | (0.046) | (0.019) | (0.017) |
| (0.031)* | (0.023)* | (0.027)* | (0.039)* | (0.036)* | (0.096)* | (0.09)* | (0.049)* | (0.032)* | (0.031)* |
| Frequency of rejections of LR test for CER: 0.068 (nominal level: 0.05) | | | | | | | | | |
| # of times in which stability is violated : 0 ; # of times in which uniqueness is violated : 0 | | | | | | | | | |

Table 1: ML estimates of the QRE-DSGE model on simulated data. NOTES: Estimates are based on M=1000 simulated samples of length T, generated from the constrained VAR (48)-(49) with k=3 lags and Gaussian disturbances with covariance matrix $\tilde{\Psi}\tilde{\Psi}'$. The values in the table are averages of the M=1000 ML estimates; the numbers in parentheses without asterisks are averages of the M=1000 estimated asymptotic standard errors obtained by replacing the unknown matrices in (32) with their consistent estimates; the numbers in parentheses with asterisks are the standard errors of the simulated distribution of M=1000 estimates. 100 samples have been discarded before starting computations and each simulated sample is initiated with 100 additional observations to get a stochastic initial state and these are then discarded. Zero values are used as starting values for the structural parameters. The LR test for CER is computed with reference to the 95 quantile from a $\chi^2$ distribution with 30-10=17 degree of freedom, where 30 is the number of estimated coefficients of the unrestricted VAR (except the covariance matrix) and 10 is the dimension of the vector of structural parameters.

VAR: $X_t = (\pi_t \ , y_t \ , i_t)'$ lags $k$=4 , 1980:3-2006:4 (including intial values), $T = 102$

Vector AR 1-5: $F(45,124) := \underset{[0.39]}{1.06}$
Vector Normality: $\chi^2(6) := \underset{[0.05]}{12.59}$

Absolute value of the two largest eigenvalues of companion matrix: 0.95 , 0.90

Cointegration rank test

| $H_0 : r \leq j$ | Trace | $p$-val |
|---|---|---|
| j=0 | 37.36 | 0.027 |
| j=1 | 20.26 | 0.048 |
| j=2 | 7.38 | 0.10 |

$$\hat{\beta} := \hat{\beta}_0 \equiv \begin{pmatrix} 1 & 0 & 0 & \underset{(0.0019)}{-0.0025} \\ 0 & -1 & 1 & \underset{(0.0073)}{-0.027} \end{pmatrix} \quad , \quad LR := \underset{[0.08]}{5.09}$$

Table 2: Diagnostic and cointegration tests on the VAR for the data. NOTES: Upper panel: Vector AR 1-5 assesses the absence of autocorrelation in the VAR disturbances against the alternative of correlation up to the fifth order; Vector normality is a multivariate test for the normality of disturbances. Middle panel: the cointegration rank test is the LR trace test (Johansen, 1996). Lower panel: estimated cointegration matrix $\beta_0$ with corresponding LR test for the over-identifying restrictions. The p-value in squared brackets is taken from a $\chi^2(2)$ distribution.

| Parameter | Estimate | Standard error |
|---|---|---|
| ML estimates | | |
| $\varpi_f$ | 0.41 | 0.19 |
| $\delta$ | 0.01 | 0.02 |
| $1-\gamma_f$ | 0.45 | 0.02 |
| $\varrho$ | 0.01 | 0.03 |
| $\lambda_r$ | 0.98 | 0.09 |
| $\lambda_y$ | 0.48 | 9.15 |
| LR test for CER: $LR := \underset{[0.00]}{73.43}$ | | |
| Parametric bootstrap $p$-value: 0.087 | | |

Table 3: Estimated structural parameters and LR test for the CER. NOTES. Upper panel: a grid search has been used for the structural parameters using, after a number of preliminary estimation trials, the range 0.15-0.45 for $\varpi_f$, the rage 0.01-0.10 for $\delta$, the range 0.35-0.48 for $1-\gamma_f$, the range 0.01-0.10 for $\varrho$ and the range 0.45-0.50 for $\lambda_y$, using 0.01 as increment. $\lambda_r$ has been estimated unrestrictedly. Lower panel: the p-value based on the asymptotic distribution is taken from a $\chi^2(15)$ distribution; the p-value obtained through the parametric bootstrap has been computed by generating M=1000 samples from the VAR with coefficients fixed at the estimates obtained from the historical sample under the CER, using the sample observations as initial values and Gaussian disturbances.