Estimation of quasi-rational DSGE models*

Luca Fanelli†

University of Bologna

October 2008

Abstract

The poor time-series performance of the class of small-scale dynamic stochastic general equilibrium (DSGE) models currently used in monetary policy, can be ascribed to the tight nature of the cross-equation restrictions these models impose on vector autoregressions (VAR) for the data, under the rational expectations hypothesis. Under these restrictions, the reduced form model solution reads as a VAR of order one and can be affected by an omitted dynamics bias, especially when quarterly time-series are involved. In this paper we argue that if the actual agents' forecast model is a finite-order VAR whose lag length is greater than one, it is possible to approximate the canonical DSGE model with a dynamic counterpart, called quasi-rational DSGE (QR-DSGE) model, whose reduced form solution has the same lag structure as the unrestricted VAR for the data. After discussing solution properties and the conditions ensuring the local identifiability of the structural parameters, we put forth a likelihood-based approach for estimating the structural parameters and testing the data adequacy of the QR-DSGE model based on an iterative switching algorithm, which exploits the asymptotic equivalence between iterated minimum-distance methods and full information maximum likelihood techniques. The analysis is briefly extended to the case of cointegrated variables. Some Monte Carlo simulations investigate the properties of the suggested approach.

Keywords: Dynamic stochastic general equilibrium model, Likelihood-based estimation, Quasi-rational expectations, VAR, Iterated minimum distance method.

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

*A previous version of this paper circulated with the title 'Estimation of a DSGE model under VAR expectations'. I wish to thank Marco Del Negro for the constructive discussion and Riccardo 'Jack' Lucchetti for helpful insights on the estimation algorithm. 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. However, despite possessing attractive theoretical properties, such as the capability of featuring potential structural sources of endogenous persistence that can potentially 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 difficulty generating sufficient endogenous persistence to match the persistence observed in the data. For this reason their empirical reliability is still an open question and misspecification remains an issue.

From the econometric point of view, DSGE models are interpreted as inherently misspecified systems (An and Schorfheide, 2007) and are usually treated as restricted but parametrically incomplete representations of the actual data: given this perspective, structural estimation and evaluation 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; Lindé, 2005), manipulating arbitrarily the shock structure of the model (Smets and Wouters, 2003, 2007), or using prior distributions with the possibility of relaxing the CER (Del Negro and Schorfheide, 2004; Del Negro et al., 2007, Del Negro and Schorfheide, 2007).[2]

When classical estimation is pursued in small scale New Keynesian DSGE models of monetary policy, the cross-equation restrictions (CER) that these models impose on the VAR for the data under the rational expectations hypothesis, are of two types: (i) a set of highly nonlinear constraints which involve the VAR coefficients and the structural parameters, in which consistent estimates of the former can be used to recover consistent estimates of the latter; (ii) a set of zero constraints which set the VAR lag order to one. Since a VAR of order one generally misrepresents the dynamic features of the data especially when quarterly time-series are involved, the estimates of the structural parameters can fail to be consistent as a result of the misspecification induced by the set of restrictions (ii). In the presence of omitted dynamics, the consequences on standard

[1] Aside from 'limited-information' techniques, the recent estimation of small-scale DSGE 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.

inferential procedures can be remarkable, see Jondeau and Le Bihan (2008).

The practise of altering/modifying the shock structure of the DSGE model to improve its time-series performance is one of the approaches which have traditionally been devised to cope with model misspecification (Del Negro and Schorfheide, 2007). Similarly, additional lags of the observable variables can be included in the baseline theoretical specification to account for real-word recognition, processing and adjustment lags as in Rudebusch (2002*a*, 2002*b*) and Fuhrer and Rudebusch (2004); see also Lindé (2005).[3]

In this paper, we argue that one relevant source of dynamic misspecification of DSGE models can be ascribed to the rational expectations hypothesis, which by construction imposes a poor lag structure to the implied reduced form equilibrium solution. Following Branch (2004), ideally VAR forecasts would correspond to rational expectations, however, in a world in which the true distribution for the economy is unknown and characterized by heterogeneous information sets, rational expectations is impossible to observe. In this setup, a VAR for the observed time-series can be regarded as 'boundedly rational' predictor which is 'in the spirit' of rational expectations. Moreover, the idea that forecasts from VARs can serve as substitutes for aggregate expectations in macroeconomic policy model is implicit in the quasi-rational expectations hypothesis, see e.g. Nerlove and Fornari (1999).[4] We point out that once the possibility that the agents in the economy are not fully rational and form their forecasts using finite-order VARs with lag length generally greater than one is taken into account, a richer dynamic structure arises as the 'natural' consequence of the departure from the rational expectations hypothesis.

We define the QR-DSGE model as a linear rational expectations model derived from the canonical DSGE model under rational expectations, whose reduced form (determinate) solution has the same lag structure as the finite-order VAR which fits the data optimally. By construction, therefore, the reduced form solution associated with the QR-DSGE model rules zero restrictions of type (ii) out. We interpret the QR-DSGE model as the dynamic counterpart of the original DSGE model which is consistent with the actual agents' expectations generating system. Thus, it implicitly accounts for the effects of omitted adjustment costs and time-to-build lags which are usually invoked in the literature to help match the data. In practice, the QR-DSGE model is obtained from the DSGE model by including a proper set of additional lags of the observable variables in the canonical structural equations. This number of additional lags, however, is not arbitrary, but corresponds to the lag order of the unrestricted VAR for the data.

[3]Lindé (2005), Section 5, p. 1146, writes: '*The additional lags in the aggregate demand equation and the monetary policy rule are required to make* ε_y *and* ε_R [the disturbances] *white noise* '.

[4]Brayton *et al.* (1997, p. 228) provide an excellent review of the concept of VAR-based expectations. Examples where VAR expectations are applied include Fuhrer and Moore (1995), Kurmann (2006) and Fanelli (2008); the seminal papers are Sargent (1979) and Campbell and Shiller (1987).

We show that the (determinate) reduced form solution associated with the QR-DSGE model can be represented as a VAR characterized by highly nonlinear coefficient restrictions of type (i) only. After discussing the conditions that ensure the generic local identifiability of the structural parameters, we discuss a procedure for maximizing the likelihood function of the system with respect to the structural parameters. More precisely, the suggested likelihood maximization algorithm exploits the asymptotic equivalence between the iterated minimum-distance procedure and the full information maximum likelihood estimator (Phillips, 1976; Hendry, 1976). Starting from the unrestricted VAR coefficient estimates, we recover the estimates of the structural parameters from the CER, by iterating the minimum distance estimator until convergence. The properties of the suggested procedure are investigated through some Monte Carlo experiments and an empirical illustration based on euro area quarterly data is provided.

The analysis is developed assuming that all time-series are generated by stationary processes, in line with the idea that DSGE models are solved by log-linearizing around a steady state. In practise, however, the steady states are usually either assumed constant (as it happens with the inflation rate and interest rates), or estimated by statistical procedures such as the Hodrick-Prescott filter. The recent literature points out that this practise may be unsatisfactory when variables can be approximated by highly persistent processes, see Dees *et al.* (2008). Moreover, treating mistakenly nonstationary as stationary processes may flaw standard inferential procedures, see Johansen (2006), Li (2007) and Fanelli (2008). We extend the analysis of the QR-DSGE model to the case in which the observed time-series can be approximated in terms of a cointegrated Vector Error Correction (VEC) system.

The structure of the paper is as follows. Section 2 presents the baseline small scale New Keynesian DSGE model and Section 3 discusses the omitted dynamics issue. Section 4 introduces the QR-DSGE model and provides results about the reduced form solution and the identifiability of the parameters. Section 5 deals with the estimation algorithm and Section 6 extends the analysis to the case of non-stationary variables. The finite sample properties of the proposed estimation algorithms are studied in Section 7 on simulated data. Some concluding remarks are provided in Section 8.

2 Model

Let $X_t = (X_{1,t}, X_{2,t}, \cdots, X_{p,t})'$ be a $p \times 1$ vector of observable variables, and assume that the New Keynesian macroeconomic system of equations can be expressed in the form

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

where, Γ_i, $i = 0, f, b$ are $p \times p$ matrices of structural parameters, c is a $p \times 1$ constant, v_t is a $p \times 1$ vector which is assumed to be adapted to the sigma-field $\mathcal{F}_t$, where $\mathcal{F}_t$ represents agents' information set at time t, and $E_t X_{t+1} \equiv E(X_{t+1} \mid \mathcal{F}_t)$. 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 the system (1) is obtained through the autoregressive specification

$$v_t = \Theta v_{t-1} + u_t \tag{2}$$

where Θ is a $p \times p$ stable matrix (i.e. with eigenvalues inside the unit circle in the complex plane) and u_t is a white noise with covariance matrix Σ_u.

Following Cho and Moreno (2006), among many others, we confine the analysis to formulations like (1)-(2) where X_t involves observable variables.[5] The matrices Γ_0, Γ_f and Γ_b are completely determined by a set of structural parameters, that we collect in the $m_s \times 1$ vector γ^s.The expression (1)-(2) does not imply that the system must have first order dynamics only: in principle, the vector X_t can always be re-defined so that (1) can be thought of as a state-space representation, or as a canonical multivariate rational expectations model as in Binder and Pesaran (2005, 2007).

According to standard practise, the v_t term in (1)-(2) is given a VAR structure with diagonal Θ for pure convenience; in general v_t provides explicit recognition that the DSGE model is not designed to capture the full extent of variation observed in the data.[6] The assumption that structural shocks may be autocorrelated is common in the literature but is also arbitrary, in the sense that it is not derived from first-principles.

Note that with $\Gamma_f = 0_{p\times p}$, the DSGE model specified in (1)-(2) collapses to a traditional simultaneous system of equations. The system (1)-(2) nests the class of small-scale (three equations) New Keynesian models typically used in monetary policy analysis, as suggested by the example below.

Example 1.

Consider the following benchmark model, consisting in the three stylized equations:

$$y_t = \varpi_f E_t y_{t+1} + (1 - \varpi_f) y_{t-1} - \delta(i_t - E_t \pi_{t+1}) + v_{1t} \tag{3}$$

$$\pi_t = \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \varrho y_t + v_{2t} \tag{4}$$

$$i_t = \lambda_r i_{t-1} + (1 - \lambda_r)(\lambda_\pi \pi_t + \lambda_y y_t) + c_3 + v_{3t} \tag{5}$$

[5] This means that we have solved problems such as how measuring potential output.

[6] 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.*'

where y_t is a measure of the output gap, π_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, (3), is a log-linearized Euler aggregate demand (IS) curve, the second equation, (4), is the New Keynesian Phillips (NKPC) curve, and the third equation, (5), 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 (3)-(5) and for the structural interpretation of the parameters (ϖ_f, δ, γ_f, γ_b, ϱ, λ_r, λ_π, λ_y).[7]

Referring to the notation used in (1), the model (3)-(5) is obtained by setting $X_t = (y_t, \pi_t, i_t)'$ and

$$\Gamma_0 = \begin{bmatrix} 1 & 0 & -\delta \\ -\varrho & 1 & 0 \\ -(1-\lambda_r)\lambda_y & -(1-\lambda_r)\lambda_\pi & 1 \end{bmatrix}, \quad \Gamma_f = \begin{bmatrix} \varpi_f & \delta & 0 \\ 0 & \gamma_f & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6}$$

$$\Gamma_b = \begin{bmatrix} (1-\varpi_f) & 0 & 0 \\ 0 & \gamma_b & 0 \\ 0 & 0 & \lambda_r \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \\ c_3 \end{bmatrix}. \blacksquare \tag{7}$$

For some parameter configurations, the rational expectations solution of the system (1)-(2) can be cast in the form

$$X_t = \Phi_1 X_{t-1} + \mu + \Psi v_t \tag{8}$$

where the $p \times p$ matrices Φ_1 and Ψ, and the $p \times 1$ vector μ are restricted as $\Phi_1 = \widetilde{\Phi}_1$, $\Psi = \widetilde{\Psi}$ and $\mu = \widetilde{\mu}$, where $\widetilde{\Phi}_1$ is a stable matrix which solves the quadratic matrix equation

$$\Gamma_f (\widetilde{\Phi}_1)^2 - \Gamma_0 \widetilde{\Phi}_1 + \Gamma_b = 0_{p \times p}, \tag{9}$$

$\widetilde{\Psi}$ is a non-singular matrix determined by

$$vec(\widetilde{\Psi}) = \left\{ [I_p \otimes (\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)] - [R' \otimes \Gamma_f] \right\}^{-1} vec(I_p) \tag{10}$$

where $vec(\cdot)$ is the column stacking operator and '$\otimes$' the Kronecker product, and $\widetilde{\mu} = (\Gamma_0 - \Gamma_f \widetilde{\Phi}_1 - \Gamma_f)^{-1} c$, see Binder and Pesaran (1995, 1997) and Uligh (1999).

[7] Following Lippi and Neri (2005), who focus on a small scale New Keynesian model for the euro area, one might further augment the system (3)-(4) by a money demand equation and derive the (possibly forward-looking) policy rule from the minimization of an intertemporal loss fuction.

Assuming that the matrices Γ_0 and $(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)$ are non-singular, it can be proved that the solution (8)-(10) is not stable (i.e. it is explosive) if the $\widetilde{\Phi}_1$ matrix has eigenvalues outside the unit circle, while the solution is not unique (i.e. there are multiple stable solutions) if $\widetilde{\Phi}_1$ is stable but the matrix $(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)^{-1}\Gamma_f$ has eigenvalues outside the unit circle, see Binder and Pesaran (1995), Section 2.3. For

Assumption 1 The matrices Γ_0 and $(\Gamma_0 - \Gamma_f \Phi_1)$ are non-singular and the matrices $(\Gamma_0 - \Gamma_f \Phi_1)^{-1}\Gamma_f$ and Θ are stable.

Hereafter the DSGE model (1)-(2) and its reduced form solution (8)-(10) under Assumption 1 will be denoted with the acronym R-DSGE to remark that the model hinges on the rational expectations hypothesis. Thus if (1)-(2) is the 'true' model and the agents in the economy are fully rational, the system (8)-(10) is the data generating process.

3 Omitted dynamics

The quadratic matrix equation (9) provides a set of CER with relate the VAR coefficients to the structural parameters of the R-DSGE model. Once the CER are deduced, the structural parameters in Γ_i, $i = 0, f, b$ can be estimated through full-information methods by maximizing one of various approximations of the likelihood function of the system, under the assumption of correct specification. For instance, one can maximize the likelihood function of the VAR (8) under the set of CER (9), see Cho and Moreno (2006). This procedure, however, can fail to deliver consistent estimates of the structural parameters because of the misspecification of the system (8) with respect to the data. With the term 'misspecification' here we mean the case of omitted dynamics in the observable variables, see Jondeau and Le Bihan (2008).

The traditional solutions which have been designed to cope with DSGE model misspecification, outside the Bayesian paradigm, include the addition of 'additional' dynamics to account for real-word recognition, processing, adjustment costs and time-to-build lags (Rudebusch 2002a, 2002b; Fuhrer and Rudebusch 2004), or the manipulation of the shock structure of the model (Smets and Wouters, 2003, 2007).

To understand the issue, observe that when in the R-DSGE model $\Theta = 0_{p\times p}$, the structural shocks follow a white noise process in (2) and the reduced form equilibrium (8) is a (stable) VAR of order one (subject to highly nonlinear CER). On the other hand, when $\Theta \neq 0_{p\times p}$, the reduced form model (8) can be equivalently written as a (stable) VAR of order two (subject to highly nonlinear CER). Indeed, by substituting (2) into (8), using some algebra and the non singularity of $\widetilde{\Psi}$, yields the expression

$$X_t = (\widetilde{\Phi}_1 + \widetilde{\Psi}\Theta\widetilde{\Psi}^{-1})X_{t-1} - \widetilde{\Psi}\Theta\widetilde{\Psi}^{-1}\widetilde{\Phi}_1 X_{t-2} + \widetilde{\mu} + \widetilde{\Psi}u_t. \tag{11}$$

By construction, the R-DSGE model (1)-(2) can give rise to a misleading representation of the probabilistic structure of the data: VAR practitioners well recognize that models with one or two lags generally represent poor dynamic characterizations of the dynamic behaviour of quarterly time-series.

If the autoregressive structure (and persistence) of the structural disturbance v_t is further enriched by considering e.g. a VAR of order two, the implied reduced form equilibrium of the R-DSGE reads as a constrained VAR of order three, and so forth. Similarly, modelling v_t as a vector autoregressive moving average process is an implicit recognition that the systematic part of model (1) omits relevant lags of X_t.

Consider an econometrician, endowed with the sample of observations $X_1, X_2, ..., X_T$ (initial values are given) on the process $\{X_t\}_{t=1}^{\infty}$. Using the available information and statistical methods, the econometrician specifies the 'best' reduced form forecast model for X_t. Suppose that the model which fits the data optimally is given by the VAR process

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \Phi_3 X_{t-3} + \Phi_4 X_{t-4} + \mu + \varepsilon_t \tag{12}$$

where Φ_i, $i = 1, ..., 4$ are $p \times p$ matrices of coefficients, μ a $p \times 1$ constant, and ε_t is a white noise process with covariance matrix $\Sigma_\varepsilon < \infty$. Driven by standard information criteria and diagnostic specification tests, the econometrician rejects the null that $\Phi_4 = 0_{p\times p}$.

Consider the following question: under which set of restrictions is the reduced form solution associated with the R-DSGE model nested within the agents' forecast model (12)? Applying the method of undetermined coefficients and setting $\varepsilon_t = \Psi u_t$, the CER between the VAR (12) and the DSGE model (1)-(2) are given by

$$\Phi_1 = (\widetilde{\Phi}_1 + \widetilde{\Psi}\Theta\widetilde{\Psi}^{-1}) \tag{13}$$

$$\Phi_2 = \widetilde{\Phi}_2 = -\widetilde{\Psi}\Theta\widetilde{\Psi}^{-1}\widetilde{\Phi}_1 \tag{14}$$

$$\Phi_3 = \widetilde{\Phi}_3 = 0_{p\times p} \quad , \quad \Phi_4 = \widetilde{\Phi}_4 = 0_{p\times p} \tag{15}$$

$$\mu = \widetilde{\mu}$$

$$\Psi = \widetilde{\Psi} \tag{16}$$

where $\widetilde{\Phi}_1$, $\widetilde{\Psi}$ and $\widetilde{\mu}$ are defined as in (9) and (10). There are two types of CER: (i) the restrictions in (13) (and in (14) when $\Theta \neq 0_{p\times p}$) which define a mapping between the VAR coefficients and the structural parameters of the R-DSGE model, which can be used to recover maximum likelihood estimates of the latter; (ii) the zero restrictions in (15) (and in (14) when $\Theta = 0_{p\times p}$) which reduce the VAR lag order from 4 to 2, in contradiction with the hypothesis $\Phi_4 \neq 0_{p\times p}$, supported by the data.

It turns out that due to the restrictions of type (ii), a VAR for X_t of lag order 4 cannot be regarded as the reduced form solution of the R-DSGE model, unless the disturbance v_t in (2) is modelled as a VAR of lag order 3, or the canonical structural equations in (1) are turned into the specification

$$\Gamma_0 X_t = \Gamma_f E_t X_{t+1} + \Gamma_b X_{t-1} + \sum_{h=2}^{4} \Omega_h X_{t-h} + c + v_t \quad , \quad v_t = u_t \sim WN(0, \Sigma_u) \tag{17}$$

for a suitable choice of the matrices of 'additional' auxiliary parameters Ω_2, Ω_3 and Ω_4. Small-scale DSGE models of the form (17) with Ω_h diagonal have been estimated in e.g. Lindè (2005) and Jondeau and Le Bihan (2008) through maximum likelihood methods. In many circumstances, the elements of Ω_h can be given structural interpretation. For instance, coming back to the Example 1, a typical dynamic approximation of the equations (3)-(5) which can be represented in the form (17) is given by

$$y_t = \varpi_f E_t y_{t+1} + (1 - \varpi_f) \sum_{h=1}^{4} \omega_h^y y_{t-h} - \delta(i_t - E_t \pi_{t+1}) + u_{1t}$$

$$\pi_t = \gamma_f E_t \pi_{t+1} + (1 - \gamma_f) \sum_{h=1}^{4} \omega_h^\pi \pi_{t-h} + \varrho y_t + u_{2t}$$

$$i_t = \sum_{h=1}^{3} \omega_h^r i_{t-h} + (1 - \sum_{h=1}^{3} \omega_h^r)(\lambda_\pi E_t \pi_{t+1} + \lambda_y y_t) + c_3 + u_{3t}$$

where the restriction $\sum_{h=1}^4 \omega_h^\pi = 1$ is consistent with the natural rate hypothesis and $\sum_{h=1}^4 \omega_h^y \leq 1$ and $\sum_{h=1}^4 \omega_h^r = 1$ are usually imposed in estimation.

In this paper, we argue that a model of the form (17) can be interpreted as the dynamic approximation of the baseline R-DSGE model (1) which is consistent, for suitable choices of the (non necessarily diagonal) matrices Ω_h, with the actual agents' forecast model in (12). More precisely, we show that under a set of regularity conditions, a restricted version of the VAR (12) can be regarded as the reduced form solution of the structural model (17), without involving restrictions of type (ii) on the VAR lag order.

4 The QR-DSGE model

Consider the VAR specification

$$X_t = \Phi_1 X_{t-1} + \cdots + \Phi_k X_{t-k} + \mu + \varepsilon_t \ , \ \varepsilon_t \sim WN(0, \Sigma_\varepsilon) \ , \ t = 1, ..., T \tag{18}$$

where Φ_j, $j = 1, ..., p$ are $p \times p$ matrices of parameters, μ is a $p \times 1$ vector of constants, ε_t is a $p \times 1$ white noise process with $p \times p$ covariance matrix Σ_ε; X_0, X_{-1}, ..., X_{-1+k} are fixed. The system

(18) is treated as the statistical model for the data, or the agents' forecast model. Conditional forecasts at time t are taken with respect to the sigma-field $\mathcal{H}_t = \sigma(X_1, ..., X_t) \subseteq \mathcal{F}_t$.

We consider the following assumptions:

Assumption 2 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 3 $\Phi_k \neq 0_{p \times p}$.

Assumption 4 The coefficients $(\Phi_1, \cdots, \Phi_k, \Sigma_\varepsilon)$ do not vary over time.

Assumption 2 rules out explosive and unit roots from the data. Actually, in some circumstances, the exclusion of unit roots may be misleading from the inferential point of view, hence we relax this assumption in Section 6. Assumption 3 implies that any model restriction which reduces the VAR lag order leads to an omitted regressors bias. Assumption 4 requires that the agents use the VAR for X_t to form their expectations over a regime where the coefficients do not drift over time. 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); hence the identification of sub-periods over which the VAR (18) fulfils Assumption 3 is crucial before the structural equations can be estimated, see Cho and Moreno (2006).

The QR-DSGE model is introduced in Defintion 1.

Definition 1

Given the VAR for the data (18) with lag length k and the R-DSGE model (1)-(2), the QR-DSGE model is defined as the multivariate linear rational expectations model

$$\Gamma_0 X_t = \Gamma_f E_t X_{t+1} + \Gamma_b X_{t-1} + \left(\sum_{j=2}^{k} \Omega_j X_{t-j} \right) \mathbb{I}_{\{k \geq 2\}} + c + v_t \tag{19}$$

$$v_t = \Theta v_{t-1} (1 - \mathbb{I}_{\{k \geq 2\}}) + u_t \tag{20}$$

where the matrices Γ_i, $i = 0, f, b$ and Θ are defined exactly as in (1)-(2), $\mathbb{I}_{\{\cdot\}}$ is the indicator function, u_t is a white noise process with covariance matrix Σ_u and the $p \times p$ matrices Ω_j may depend on Γ_i, $i = 0, f, b$ and Φ_h, $h = 1, ..., k$, for $j = 2, ..., k$.

The main difference between the QR-DSGE model (19)-(20) and the R-DSGE model (1)-(2) is that the former involves, by construction, a number of k additional auxiliary lags of X_t (when

$k \geq 2$), where k is determined from the data. Observe that with $k = 1$ the QR-DSGE model (19)-(20) collapses to the canonical R-DSGE model (1)-(2). In this setup, the parameters in the Ω_j matrices are not intended to capture micro-founded propagation mechanisms but are necessary to match the lag structure of the agents' expectations generating system.

The proposition that follows discusses the solution of the QR-DSGE model.

Proposition 1

If Assumptions 1-3 are satisfied, the determinate reduced form solution of the QR-DSGE model (19)-(20), if it exists, is given by the VAR in (18) where $\varepsilon_t = \Psi u_t$, $\Phi_j = \widetilde{\Phi}_j$, $j = 1, ..., k$, $\Psi = \widetilde{\Psi}$, $\mu = \widetilde{\mu}$, and the matrices $\widetilde{\Phi}_j$, $\widetilde{\Psi}$ and $\widetilde{\mu}$ fulfil the following set of restrictions:

$$\begin{aligned}
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)\widetilde{\Phi}_1 &= \Gamma_f \widetilde{\Phi}_2 + \Gamma_b && (21)\\
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)\widetilde{\Phi}_2 &= \Gamma_f \widetilde{\Phi}_3 + \Omega_2 && (22)\\
&\vdots \\
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)\widetilde{\Phi}_{k-1} &= \Gamma_f \widetilde{\Phi}_k + \Omega_{k-1} \\
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)\widetilde{\Phi}_k &= \Omega_k \\
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1 - \Gamma_f)\widetilde{\mu} &= c && (23)\\
(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1)\widetilde{\Psi} &= I_p. && (24)
\end{aligned}$$

Proof: See Appendix A.

Observe that (24) implies that the VAR covariance matrix is restricted as $\Sigma_\varepsilon = \widetilde{\Psi}\Sigma_u\widetilde{\Psi}'$. Moreover, when $k = 1$ the reduced form solution of the system amounts to the constrained VAR of order one in (8)-(10). A corollary of Proposition 1 is related to the 'optimal' choice of the matrices Ω_j in (19)-(20).

Corollary 1

If the matrices Ω_j in the QR-DSGE model of Definition 1 are defined as

$$\begin{aligned}
\Omega_j &= (\Gamma_0 - \Gamma_f \Phi_1)\Phi_j - \Gamma_f \Phi_{j+1} \quad , j = 2, ..., k-1 && (25)\\
\Omega_k &= (\Gamma_0 - \Gamma_f \Phi_1)\Phi_k, && (26)
\end{aligned}$$

the CER collapse, for $\Phi_1 = \widetilde{\Phi}_1$, to (21) and (23)-(24), and the matrices Φ_2, $\Phi_3, ..., \Phi_k$ are unrestricted.

Corollary 1 suggests that the QR-DSGE model in which the auxiliary matrices Ω_j are specified as in (25)-(26) is consistent with a determinate reduced form solution, if it exists, having the same lag structure as the VAR for the data and involving the 'minimal' set of nonlinear CER. In this case, the auxiliary parameters depend on the structural parameters and VAR coefficient alone and not on additional parameters, albeit it is not clear which is the interpretation attached to the choice (25)-(26).

Let $\Gamma = (\Gamma^s : \Omega)$, $\Gamma^s = (\Gamma_0 : \Gamma_f : \Gamma_b)$, $\Omega = (\Omega_2 : ... : \Omega_k)$ be the $p \times (3p + k - 1)$ matrix summarizing all parameters (the structural in Γ^s and the auxiliary in Ω) of the QR-DSGE model, except the parameters of the covariance matrix. The link between Γ and the $m \times 1$ vector $\gamma = (\gamma^{s\prime}, \omega')'$, where ω is the $m_\omega \times 1$ vector containing the unrestricted (free) elements of Ω ($m = m_s + m_\omega$),[8] can be specified as

$$vec(\Gamma) = q(\gamma) \tag{27}$$

where $q(\cdot)$ is a twice differentiable vector function such that the $p(3p + k - 1) \times m$ Jacobian matrix $Q_\gamma = \partial q(\gamma)/\partial \gamma'$ has full column rank. If the relation between Γ and γ is linear, the expression (27) collapses to $vec(\Gamma) = Q\gamma$, where $Q_\gamma = Q$ does not depend on γ and contains zeros and ones only. The next proposition deals with the local identifiability of the parameters of the QR-DSGE model.

Proposition 2

Consider the VAR (18) and assume that $k \geq 2$. (i) Necessary condition for the local identifiability of the parameters γ of the QR-DSGE model is that $m \leq p^2(k+1)$; if $m < p^2(k+1)$ there are $p^2(k+1) - m$ over-identifying restrictions. (ii) If the order condition and the Assumptions 1-3 hold, the VAR coefficients in (21)-(23) can be uniquely expressed as function of γ in a neighborhood of true parameter values.

Proof: Appendix A.

5 Estimation

According to Proposition 2, the VAR (18) subject to the restrictions (21)-(24) can be written in the form

$$X_t = \widetilde{\Phi}(\gamma) X^*_{t-1} + \varepsilon_t \tag{28}$$

where $\widetilde{\Phi}(\gamma) = [\widetilde{\Phi}_1(\gamma) : \cdots : \widetilde{\Phi}_k(\gamma) : \widetilde{\mu}(\gamma)]$, $X^*_{t-1} = (X'_{t-1}, ..., X'_{t-k}, 1)'$, and $\varepsilon_t = \widetilde{\Psi}(\gamma) u_t$. The notation used in (28) remarks the dependence of the restricted VAR coefficients on the structural parameters γ; in particular, $\widetilde{\Phi}_j(\gamma) = \widetilde{\Phi}_j$, $j = 1, ..., k$, $\widetilde{\mu}(\gamma) = \widetilde{\mu}$, and $\widetilde{\Psi}(\gamma) = \widetilde{\Psi}$, where the $\widetilde{\Phi}_j$, $\widetilde{\mu}$ and $\widetilde{\Psi}$ are determined in (21)-(24).

[8] Note that $m_\omega = 0$ in the specification (25)-(26).

Since it is difficult to get a closed form (analytic) expression in which the VAR coefficients are expressed as function of γ, one possible method is the minimum-distance approach which asymptotically leads to a quasi-maximum likelihood estimator for γ, see Phillips (1976).

Exploiting this intuition, a possible way to get maximum likelihood estimates of γ, is to rewrite the CER (21)-(23) in the form $\widetilde{\Phi} = F(\widetilde{\Phi}, \gamma)$, and define Newton-like iterations of the form

$$\widetilde{\Phi}^{(g+1)} \simeq F(\widetilde{\Phi}^{(g)}, \gamma^{(g)}) \ , \quad g = 0, 1, ... \tag{29}$$

A 'natural' starting point of the algorithm is given by the choice $\widetilde{\Phi}^{(0)} = \widehat{\Phi}$, $\widehat{\Phi}$ being the unrestricted least squares (maximum likelihood) estimate of the VAR coefficients.[9] Assume that $\gamma^{(0)} = \gamma$; at $g = 1$, if the matrix $(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1^{(0)})$ is non singular, the iteration (29) corresponds to the set of explicit form constraints

$$\begin{aligned}
\widetilde{\Phi}_{1,\gamma}^{(1)} &\simeq (\Gamma_0 - \Gamma_f \ \widetilde{\Phi}_1^{(0)})^{-1}(\Gamma_f \ \widetilde{\Phi}_2^{(0)} + \Gamma_b) && (30)\\
\widetilde{\Phi}_{2,\gamma}^{(1)} &\simeq (\Gamma_0 - \Gamma_f \ \widetilde{\Phi}_1^{(0)})^{-1}(\Gamma_f \ \widetilde{\Phi}_2^{(0)} + \Omega_2) \\
&\ \ \vdots \\
\widetilde{\Phi}_{k,\gamma}^{(1)} &\simeq (\Gamma_0 - \Gamma_f \ \widetilde{\Phi}_1^{(0)})^{-1}\Omega_k && (31)
\end{aligned}$$

which express the reduced form VAR coefficients as function of the (unknown) elements in γ. The BFGS method (Fletcher, 1987) can be then used to maximize the VAR likelihood under the restrictions (30)-(31) with respect to γ, obtaining an initial estimate of γ, $\widehat{\gamma} = \widehat{\gamma}^{(0)}$, and the corresponding value of the likelihood. The elements in $\widehat{\gamma}^{(0)}$ are then used back to update the reduced form VAR coefficients

$$\begin{aligned}
\widetilde{\Phi}_1^{(1)} &= \widetilde{\Phi}_{1,\widehat{\gamma}}^{(1)} \simeq (\widehat{\Gamma}_0^{(0)} - \widehat{\Gamma}_f^{(0)} \ \widetilde{\Phi}_1^{(0)})^{-1}(\widehat{\Gamma}_f^{(0)} \ \widetilde{\Phi}_2^{(0)} + \widehat{\Gamma}_b^{(0)}) \\
\widetilde{\Phi}_2^{(1)} &= \widetilde{\Phi}_{2,\widehat{\gamma}}^{(1)} \simeq (\widehat{\Gamma}_0^{(0)} - \widehat{\Gamma}_f^{(0)}\widetilde{\Phi}_1^{(0)})^{-1}(\widehat{\Gamma}_f^{(0)} \ \widetilde{\Phi}_2^{(0)} + \widehat{\Omega}_2^{(0)}) \\
&\ \ \vdots \\
\widetilde{\Phi}_k^{(1)} &= \widetilde{\Phi}_{k,\widehat{\gamma}}^{(1)} \simeq (\widehat{\Gamma}_0^{(0)} - \widehat{\Gamma}_f^{(0)} \ \widetilde{\Phi}_1^{(0)})^{-1}\widehat{\Omega}_k^{(0)}
\end{aligned}$$

which are then used at the iteration $g = 2$. At $g = 2$, assuming that the matrix $(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1^{(1)})$ is invertible,[10] the expression (29) amounts to the updated set of constraints

[9] Clearly, $\widehat{\Phi}$ will read as a quasi-maximum likelihood estimator if the VAR disturbances are not Gaussian.

[10] The matrix $(\Gamma_0 - \Gamma_f \widetilde{\Phi}_1^{(0)})$ can also be kept fixed in the expressions above across iterations.

$$
\begin{aligned}
\widetilde{\Phi}^{(2)}_{1,\gamma} &\simeq (\Gamma_0 - \Gamma_f\ \widetilde{\Phi}^{(1)}_1)^{-1}(\Gamma_f\ \widetilde{\Phi}^{(1)}_2 + \Gamma_b) \\
\widetilde{\Phi}^{(2)}_{2,\gamma} &\simeq (\Gamma_0 - \Gamma_f\ \widetilde{\Phi}^{(1)}_1)^{-1}(\Gamma_f\ \widetilde{\Phi}^{(1)}_2 + \Omega_2) \\
&\vdots \\
\widetilde{\Phi}^{(2)}_{k,\gamma} &\simeq (\Gamma_0 - \Gamma_f\ \widetilde{\Phi}^{(1)}_1)^{-1}\Omega_k
\end{aligned}
$$

which can be again exploited to get the estimate $\widehat{\gamma} = \widehat{\gamma}^{(1)}$ through the BFGS method and a the updated value of the likelihood. The elements in $\widehat{\gamma}^{(1)}$ are used to update the reduced form VAR coefficients according to

$$
\begin{aligned}
\widetilde{\Phi}^{(2)}_1 &= \widetilde{\Phi}^{(2)}_{1,\widehat{\gamma}} \simeq (\widehat{\Gamma}^{(1)}_0 - \widehat{\Gamma}^{(1)}_f\ \widetilde{\Phi}^{(1)}_1)^{-1}(\widehat{\Gamma}^{(1)}_f\ \widetilde{\Phi}^{(1)}_2 + \widehat{\Gamma}^{(1)}_b) \\
\widetilde{\Phi}^{(2)}_2 &= \widetilde{\Phi}^{(2)}_{2,\widehat{\gamma}} \simeq (\widehat{\Gamma}^{(1)}_0 - \widehat{\Gamma}^{(1)}_f \widetilde{\Phi}^{(1)}_1)^{-1}(\widehat{\Gamma}^{(1)}_f\ \widetilde{\Phi}^{(1)}_2 + \widehat{\Omega}^{(1)}_2) \\
&\vdots \\
\widetilde{\Phi}^{(2)}_k &= \widetilde{\Phi}^{(2)}_{k,\widehat{\gamma}} \simeq (\widehat{\Gamma}^{(1)}_0 - \widehat{\Gamma}^{(1)}_f\ \widetilde{\Phi}^{(1)}_1)^{-1}\widehat{\Omega}^{(1)}_k,
\end{aligned}
$$

and so forth. The procedure is iterated for $g = 3, 4, ...,$ until convergence.[11]

Proposition 3

Given the constrained VAR (28), let $\widehat{\gamma}$ be the estimator of the structural parameters resulting from the algorithm described above. Under Assumption 2-4, $T^{1/2}(\widehat{\gamma} - \gamma)$ is asymptotically Gaussian with covariance matrix given by the expression

$$
V_\gamma = \overline{J}'_{\gamma,\phi} V_{\widetilde{\Phi}} \overline{J}_{\gamma,\phi} \tag{32}
$$

where $V_{\widetilde{\Phi}} = \Sigma_\varepsilon(\gamma) \otimes \Upsilon^{-1}_{xx}$ is the covariance matrix of the constrained reduced form coefficients, $\Sigma_\varepsilon(\gamma) = E[(X_t - \widetilde{\Phi}(\gamma)X^*_{t-1})(X_t - \widetilde{\Phi}(\gamma)X^*_{t-1})']$, $\Upsilon_{xx} = E(X^*_t X^{*\prime}_t)$, $\overline{J}_{\gamma,\phi} = J_{\gamma,\phi}(J'_{\gamma,\phi}J_{\gamma,\phi})^{-1}$ and the Jacobian $J_{\gamma,\phi}$ is defined in equation (53) in the Appendix.

Proof: Appendix A.

A consistent estimate of the covariance matrix V_γ can be obtained by replacing $V_{\widetilde{\Phi}}$ with $\Sigma_\varepsilon(\widehat{\gamma}) \otimes \widehat{\Upsilon}^{-1}_{xx}$, $\widehat{\Sigma}_\varepsilon(\widehat{\gamma}) = T^{-1}\sum_{t=1}^{T}(X_t - \widetilde{\Phi}(\widehat{\gamma})X^*_{t-1})(X_t - \widetilde{\Phi}(\widehat{\gamma})X^*_{t-1})'$, $\widehat{\Upsilon}_{xx} = T^{-1}\sum_{t=1}^{T} X^*_t X^{*\prime}_t$ and replacing the unknown parameters in the Jacobian $J_{\gamma,\phi}$ by their consistent estimates.

[11] Observe that the VAR coefficients $\widetilde{\Phi}^{(g)}_i$ can be computed through quasi-Newton algorithms, instead of using the explicit expressions above, possibly avoiding the inversion of the matrix $(\widehat{\Gamma}^{(g)}_0 - \widehat{\Gamma}^{(g)}_f \widetilde{\Phi}^{(g)}_1)$. This issue is the object of further investigation.

6 Non-stationary variables

The issue of non-stationary time-series is usually ignored in small-scale DSGE models, despite the time series pattern of variables such as the inflation rate and the interest rates display a persistence which hardly can be removed by simply subtracting constants representing deterministic steady state levels.

In this section we briefly extend the econometric analysis of the QR-DSGE model to the case where Assumption 2 is replaced by:

Assumption 2' The roots, s, of $\det[\Phi(s)] = 0$ are such that $\mid s \mid\geq 1$; in particular, there are exactly $p-r$ roots at $s = 1$, where $0 < r < p$.

Under Assumption 2', the VAR in (18) can be represented in Vector Error Correction (VEC) form

$$\Delta X_t = \alpha\beta' X_{t-1} + \Xi W_{t-1} + \mu + \varepsilon_t \ , \ \varepsilon_t \sim WN(0,\Sigma_\varepsilon) \ , \ t = 1, ..., T \tag{33}$$

where α and β are $p \times r$ matrices of full rank r, respectively, $\Xi = [\Xi_1 : \Xi_2 : ... : \Xi_{k-1}]$ and $W_{t-1} = (\Delta X'_{t-1}, \Delta X'_{t-2}, ..., \Delta X'_{t-k+1})'$, see Johansen (1996). For fixed values of $\beta = \beta^0$, where β^0 is any identifiable version of the cointegration relations, the elements in $\beta'_0 X_t$ corresponds to the linear combinations of the variables in X_t that can be approximated as stationary processes. If, for instance, the output gap is the only stationary variable in $X_t = (y_t, \pi_t, i_t)'$, $r = 1$ and $\beta_0 = (1, 0, 0)'$; if also a Fisher parity relation between i_t and π_t holds making the ex-post real interest rate 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, and β has been fixed at β_0 as in (**??**), a convenient (triangular) representation of the VEC (33) is obtained by defining the $p \times 1$ vector

$$Y_t = \begin{pmatrix} \beta'_0 X_t \\ \tau' \Delta X_t \end{pmatrix} \begin{matrix} r \times 1 \\ (p-r) \times 1 \end{matrix} \tag{34}$$

where τ is a $p \times (p-r)$ selection matrix such that $\det(\tau'\beta_{0\perp}) \neq 0$, $\beta_{0\perp}$ being the orthogonal complement of β_0 (Johansen, 1996). In the two examples above, one has $Y_t = (y_t, \Delta\pi_t, \Delta i_t)'$ with $\tau = (e_2 : e_3)$, $e'_2 = (0, 1, 0)$, $e'_3 = (0, 0, 1)$, and $Y_t = (y_t, i_t - \pi_t, \Delta\pi_t)'$ with $\tau = e_2$, respectively.

The vector Y_t in (34) admits the stable VAR representation

$$Y_t = \Lambda_1 Y_{t-1} + \Lambda_2 Y_{t-2} + ... + \Lambda_k Y_{t-k} + \mu^Y + \varepsilon_t^Y \tag{35}$$

where the matrices Λ_j, $j = 1, ..., k$ and the vector μ^Y are suitable function of the elements in α, Ξ, μ, and $\varepsilon_t^Y = (\beta,\tau)'\varepsilon_t$, see Mellader *et al.* (1992) and Paruolo (2003), Theorem 2.[12] Note,

[12] If the constant μ in the VEC (33) 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 (34), and μ^Y may be possibly zero in the VAR (35).

in particular, that the Λ_k matrix in (35) is constrained as

$$\Lambda_k = \Lambda_k^* \equiv \left[\Lambda_{1k} \ : \ 0_{p\times(p-r)}\right]. \tag{36}$$

For the VAR in (35) to read as the reduced-form solution of the QR-DSGE model, the structural equations can be reparameterized in terms of Y_t. Consider, for $k \geq 2$, the following formulation of the QR-DSGE model

$$\Gamma_0^y Y_t = \Gamma_f^y E_t Y_{t+1} + \Gamma_b^y Y_{t-1} + \sum_{j=2}^{k} \Omega_j^y Y_{t-j} + c^y + u_t^y \tag{37}$$

which involves, by construction, only stationary variables. The superscript 'y' over the matrices of parameters remarks that the specification (37) is obtained as a reparameterization or reformulation of the original equations of the QR-DSGE model.

The example below shows how a specification of the form (37) can be obtained in practise.

Example 2

Turning on the Example 1, assume that X_t is non-stationary, and there is one common stochastic trend in the system, hence $r = 2$ in (33). Assume further that $\beta = \beta_0 = (\beta_{01} : \beta_{02})$, where $\beta_{01} = (1, 0, 0)'$, and $\beta_{02} = (0, -1, 1)'$, meaning that the output gap, y_t, and the ex-post real interest rate, $i_t - \pi_t$, are stationary; from (34), $Y_t = (y_t, i_t - \pi_t, \Delta\pi_t)'$. Using the variables in Y_t, and imposing the restriction $\gamma_f + \gamma_b = 1$ on (4), the equations (3), (4) and (5) can be reparameterized in the form

$$y_t = \varpi_f E_t y_{t+1} + (1 - \varpi_f) y_{t-1} - \delta(i_t - \pi_t) - \delta E_t \Delta\pi_{t+1} + v_{1t} \tag{38}$$

$$\Delta\pi_t = \frac{(1-\gamma_b)}{\gamma_b} E_t \Delta\pi_{t+1} + \frac{\varrho}{\gamma_b} y_t + \frac{1}{\gamma_b} v_{2t} \tag{39}$$

$$\Delta i_t = (\lambda_r - 1)(i_{t-1} - \lambda_\pi \pi_{t-1}) + (1 - \lambda_r)\lambda_\pi \Delta\pi_t + (1 - \lambda_r)\lambda_y y_t + c_3 + v_{3t} \tag{40}$$

and the QR-DSGE specification can be obtained accordingly. It can be recognized that for $\lambda_\pi = 1$, all three equations involve variables (or expectations of variables) in Y_t.$\blacksquare$

By replacing the VAR in (18) with the VAR in (35), and the QR-DSGE model (19)-(20) with (37), the propositions 1 and 2 can be still applied so that the implied set of CER are given by

$$
\begin{aligned}
(\Gamma_0^y - \Gamma_f^y \widetilde{\Lambda}_1)\widetilde{\Lambda}_1 &= \Gamma_f^y \widetilde{\Lambda}_2 + \Gamma_b^y \qquad (41)\\
(\Gamma_0^y - \Gamma_f^y \widetilde{\Lambda}_1)\widetilde{\Lambda}_2 &= \Gamma_f^y \widetilde{\Lambda}_3 + \Omega_2^y \\
&\vdots \\
(\Gamma_0^y - \Gamma_f^y \widetilde{\Lambda}_1)\widetilde{\Lambda}_{k-1} &= \Gamma_f^y \widetilde{\Lambda}_k + \Omega_{k-1}^y \\
(\Gamma_0^y - \Gamma_f^y \widetilde{\Lambda}_1)\widetilde{\mu}^y &= \Gamma_f^y \widetilde{\mu}^y + c \qquad (42)\\
(\Gamma_0^y - \Gamma_f^y \widetilde{\Lambda}_1)\widetilde{\Psi}^y &= I_p
\end{aligned}
$$

where $\widetilde{\Lambda}_j$, $j = 1, ..., k$, $\widetilde{\mu}^y$ are the restricted counterparts of the coefficients in (35)-(36) and $\varepsilon_t^y = \widetilde{\Psi}^y u_t^y$.

The estimation of the cointegrated QR-DSGE model can be carried out as follows. When β_0 is not known, once the maximum likelihood estimate $\widehat{\beta}_0$ has been obtained from the VEC (33), it can be treated as the 'true' value of β_0 in (34) because of the super-consistency result (Johansen, 1996). The algorithm described in Section 5 can be applied to the VAR (35)-(36) under the CER (41)-(42).

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.

The reference reduced form is a VAR of the form (18) of lag order $k = 3$ and with Gaussian disturbances. The baseline structural model is given by the three equations of Example 1, in which we replace the term $(1 - \lambda_r)\lambda_\pi \pi_t$ in the policy rule (5) with $(1 - \lambda_r)\lambda_\pi E_t \pi_{t+1}$, obtaining a forward-looking Taylor type rule.

For computational convenience, the matrices Ω_2 and Ω_3 are specified as $\Omega_2 = \omega_2 I_3$, and $\Omega_3 = \omega_3 I_3$, with ω_2 and ω_3 scalars and the constant c is set to zero. The covariance matrix of the disturbance u_t is taken as $\Sigma_u = I_3$, so that the covariance matrix of $\varepsilon_t = \widetilde{\Psi} u_t$ is given by $\Sigma_\varepsilon = \widetilde{\Psi}\widetilde{\Psi}'$.

The values of the structural parameters of the QR-DSGE model analyzed in this experiment are reported in Table 1; $M = 1000$ samples of length $T = 200$ and 500 respectively, have been generated from the reduced form (18) subject to the CER (21)-(23).[13] We use a relatively high persistent VAR (the absolute value of the two largest roots of the restricted companion matrix

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

is equal to 0.95) to mimic situations that may occur in practise. The maximum likelihood estimates are presented in Table 1. The table also reports the rejection frequency (empirical level) of the LR test for the CER computed using the 5% nominal critical values.

The results of this experiment show that even with a sample of $T = 200$ observations, asymptotic standard errors may mistakenly lead to deem part of the structural parameters (δ and λ_y) insignificant. Note, in particular, that the parameter δ captures the monetary policy channel, i.e. the contemporaneous output gap dependence on the ex ante real rate of interest, whereas λ_y measures the long run response of the Central Bank to output gap fluctuations. Since a sample of $T = 200$ quarterly observations without policy regimes changes can be hardly expected to be observed in practise, the results in Table 1 support Cho and Moreno's (2006) small sample approach to the estimation and testing of New Keynesian macro models. Moreover, the empirical level of the LR test for the CER appears remarkably higher than the nominal level also with $T = 500$ observations, suggesting that overrejection is an issue that can be addressed by the use of simulation-based methods.

8 Concluding remarks

In this paper we have argued that the poor time-series performance of small-scale DSGE models can be ascribed to the inadequacy of the underlying rational expectations paradigm. We have relaxed the rational expectations hypothesis in favour of a particular formulation of the quasi-rational expectations hypothesis, in which the agents use VARs as their statistical model for the data to form expectations.

The QR-DSGE model reads as a dynamic approximation of the canonical DSGE model and has a reduced form solution which is consistent with the agents' statistical model. By construction it involves a richer dynamic structure compared to the system counterpart based on the rational expectations hypothesis. A likelihood-based estimation algorithm for the QR-DSGE model has been provided.

A Monte Carlo experiment has shown that inference may be imprecise in samples of the sizes typically available to macroeconomists due to the highly nonlinear nature of the restrictions. It has been also shown how the analysis can extended to the case of non-stationary, cointegrated processes.

A Appendix

Proof of Proposition 1

If $k = 1$ the proof is straightforward, see Binder and Pesaran (1995). If $k \geq 2$, write the QR-DSGE model in the form

$$\Gamma_0^* X_t^* = \Gamma_b^* X_{t-1}^* + \Gamma_f^* E_t X_{t+1}^* + c^* + u_t^* \tag{43}$$

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

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

$$\Gamma_b^* = \begin{bmatrix} \Gamma_b & \Omega_2 & \cdots & \Omega_k \\ I_p & 0_{p\times p} & \cdots & 0_{p\times p} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{p\times p} & 0_{p\times p} & I_p & 0_{p\times p} \end{bmatrix}.$$

Write also the VAR (18) in first-order companion form

$$X_t^* = \Phi^* X_{t-1}^* + \mu^* + \varepsilon_t^* \tag{44}$$

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

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

where Φ^* is stable by Assumption 1. Provided that Γ_0^* and $(\Gamma_0^* - \Gamma_f^* \widetilde{\Phi}^*)$ are non-singular, if a unique and stable solution of the system (43) exists, it takes the form (44) with $\varepsilon_t^* = \Psi^* u_t^*$ and coefficients subject to the restrictions $\Phi^* = \widetilde{\Phi}$, $\Psi = \widetilde{\Psi}$, $\mu^* = \widetilde{\mu}^*$, where $\widetilde{\Phi}^*$, $\widetilde{\Psi}^*$and $\widetilde{\mu}^*$ are given by

$$\Gamma_f^* (\widetilde{\Phi}^*)^2 - \Gamma_0^* \widetilde{\Phi}^* - \Gamma_b^* = 0_{pk\times pk} \tag{45}$$

and

$$\mu^* = (\Gamma_0^* - \Gamma_f^* \widetilde{\Phi}^* - \Gamma_f^*)^{-1} c^* \tag{46}$$

$$\Psi^* = (\Gamma_0^* - \Gamma_f^* \widetilde{\Phi}^*)^{-1}. \tag{47}$$

Provided that the $\widetilde{\Phi}^*$ matrix solving (45) is stable, the stability of the matrix $(\Gamma_0^* - \Gamma_f^* \widetilde{\Phi}^*)^{-1}$ Γ_f^* is sufficient for the solution to be unique and stable. Γ_0^* is non-singular by Assumption 1.

Moreover, the matrix

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

is non-singular if the sub-matrix in the upper-left corner is non-singular, condition guaranteed by Assumption 1. Finally, by using inversion formulas for partitioned matrix, one gets

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

and this matrix is stable if the sub-matrix in the upper-left corner is stable, condition guaranteed by Assumption 1. It turns out that the VAR (44) subject to the restrictions (45)-(47) is the determinate solution to (43). Using the definition of the variables, the expressions in (45)-(47) correspond to (21)-(24). Note that Assumption 3 prevents that $\Omega_k = 0_{p\times p}$. ■

Proof of Proposition 2

Write the CER (21)-(23) compactly as

$$\Gamma_0(\widetilde{\Phi} : \widetilde{\mu}) - \Gamma_f \widetilde{\Phi}_1 (\widetilde{\Phi} : \widetilde{\mu}) - \Gamma_f (\widetilde{\Phi} : \mu) K - \Gamma_b^0 = 0_{p\times(pk+1)} \tag{48}$$

where K is a $p(k+1) \times p(k+1)$ selection matrix of suitable dimensions such that $(\widetilde{\Phi} : \widetilde{\mu})K = (0_{p\times p} : \widetilde{\Phi}_2 : \cdots : \widetilde{\Phi}_k : \widetilde{\mu})$, and $\Gamma_b^0 = (\Gamma_b : \Omega)$, $\Omega = (\Omega_2 : ... : \Omega_k)$. Let $\widetilde{\phi} = vec(\widetilde{\Phi} : \widetilde{\mu})$, $\widetilde{\phi}_j = vec(\widetilde{\Phi}_j)$, $j = 1, ..., k$, where $\widetilde{\phi} = (\widetilde{\phi}_1', ..., \widetilde{\phi}_k', \widetilde{\mu}')'$. Applying the vec operator to both sides of (48) yields the relation

$$(I_{p(pk+1)} \otimes \Gamma_0)\widetilde{\phi} - (I_{p(pk+1)} \otimes \Gamma_f \widetilde{\Phi}_1)\widetilde{\phi} - (K' \otimes \Gamma_f)\widetilde{\phi} - vec(\Gamma_b^0) = 0_{a\times 1} \tag{49}$$

where $a = p(pk+1)$. The relation (49) defines the vector function

$$f(\widetilde{\phi}, \gamma) = 0_{a\times 1}$$

where $f : \mathbb{S} \to \mathbb{R}^a$ is continuous differentiable on the set $\mathbb{S}$ of $\mathbb{R}^{a+m}$, and γ has been defined in (27). By the implicit function theorem, the elements in $\widetilde{\phi}$ can be uniquely expressed as function of γ in a neighborhood $N(\gamma_0) \subset \mathbb{R}^h$ of the true parameter value, γ_0, if the $a \times a$ Jacobian matrix

$$D = \frac{\partial f(\widetilde{\phi}, \gamma)}{\partial \widetilde{\phi}'} \tag{50}$$

is non-singular at $(\widetilde{\phi}_0,\, \gamma_0)$; in particular, $\widetilde{\phi} = g(\gamma)$ for all and γ in $N(\gamma_0)$, where g is a differentiable function $g : N(\gamma_0) \rightarrow \mathbb{R}^a$. Thus, the order condition $m \leq a$ is necessary for the constrained reduced form coefficients not to exceed the unrestricted VAR coefficients; if $m < a$, there are $a - m$ over-identifying restrictions in the system. This proves (i).

To compute the Jacobian matrix (50) we apply the vec operator to each block of restrictions in (21)-(24) and decompose D in blocks of dimension $p^2 \times p^2$, respectively. For instance, taking the vec of (21) and deriving the resulting expression with respect to the vectors $\widetilde{\phi}_l$, $l = 1, ..., k$ and $\widetilde{\mu}$, one gets

$$D_{1,1} = (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f)[(\widetilde{\Phi}_1' \otimes I_p) + (I_p \otimes \widetilde{\Phi}_1)];$$

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

where it is implicitly assumed that the derivatives above are evaluated at the 'true' parameter values $(\widetilde{\phi}_0,\, \gamma_0)$. Similarly, taking the vec of (22) and deriving the resulting expression with respect to the vectors $\widetilde{\phi}_l$, $l = 1, ..., k$ and $\widetilde{\mu}$ yields:

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

and so forth. The resulting D matrix reads as

$$D = \begin{bmatrix} D_{1,1} & -(I_p \otimes \Gamma_f) & 0_{p^2 \times p^2} & 0_{p^2 \times p^2} & \cdots & 0_{p^2 \times p} \\ -(\widetilde{\Phi}_2' \otimes \Gamma_f) & D_{2,2} & -(I_p \otimes \Gamma_f) & 0_{p^2 \times p^2} & \cdots & 0_{p^2 \times p} \\ -(\widetilde{\Phi}_3' \otimes \Gamma_f) & 0_{p^2 \times p^2} & D_{3,3} & -(I_p \otimes \Gamma_f) & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -(\widetilde{\Phi}_k' \otimes \Gamma_f) & 0_{p^2 \times p^2} & 0_{p^2 \times p^2} & \cdots & D_{k,k} & 0_{p^2 \times p} \\ -(\widetilde{\mu}' \otimes \Gamma_f) & 0_{p \times p^2} & 0_{p \times p^2} & \cdots & 0_{p \times p^2} & D_{k+1,k+1} \end{bmatrix} \tag{51}$$

where $D_{l,l} = (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f \widetilde{\Phi}_1)$, $l = 2, ..., k$, and $D_{k+1,k+1} = (\Gamma_0 - \Gamma_f \widetilde{\Phi}_1 - \Gamma_f)$. It can be noticed that by construction, each block of p^2 columns of D can not be obtained as linear combination of the other blocks of p^2 columns. The $D_{l,l}$, $l = 2, ..., k$ matrices on the main diagonal are non-singular by Assumption 1. Moreover, since $D_{k+1,k+1} = (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1) - \Gamma_f = (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)[I_p - (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)^{-1}\Gamma_f]$, also $D_{k+1,k+1}$ is non-singular by Assumption 1. Finally,

$$\begin{aligned} D_{1,1} &= (I_p \otimes \Gamma_0) - (I_p \otimes \Gamma_f)[(\widetilde{\Phi}_1' \otimes I_p) + (I_p \otimes \widetilde{\Phi}_1)] \\ &= [I_p \otimes (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)] - (\widetilde{\Phi}_1' \otimes \Gamma_f) \\ &= [I_p \otimes (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)][I_{p^2} - (I_p \otimes (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)^{-1})(\widetilde{\Phi}_1' \otimes \Gamma_f)] \\ &= [I_p \otimes (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)][I_{p^2} - (\widetilde{\Phi}_1' \otimes (\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)^{-1}\Gamma_f)] \end{aligned}$$

where the second matrix on the right is non singular as both $\widetilde{\Phi}_1$ and $(\Gamma_0 - \Gamma_f\widetilde{\Phi}_1)^{-1}\Gamma_f$ are stable. We have therefore proved that each block of p^2 columns in (51) has rank p^2; it turns out that $rank(D) = kp^2 + p = a$.■

Proof of Proposition 3

Given $\widehat{\gamma}$, let $\widetilde{\Phi}(\widehat{\gamma})$ be the estimator of the restricted VAR coefficients. Under Assumption 2 one has

$$T^{1/2}vec\left[\widetilde{\Phi}(\widehat{\gamma}) - \widetilde{\Phi}(\gamma)\right] \underset{T\to\infty}{\to} N\left(0_{a\times 1}\ ,\ V_{\widetilde{\Phi}}\right)$$

where

$$V_{\widetilde{\Phi}} = \Sigma_\varepsilon(\gamma) \otimes \Upsilon_{xx}^{-1}\ ,\ \Sigma_\varepsilon(\gamma) = E[(X_t - \widetilde{\Phi}(\gamma)X_{t-1}^*)(X_t - \widetilde{\Phi}(\gamma)X_{t-1}^*)']\ \ ,\ \Upsilon_{xx} = E(X_t^* X_t^{*\prime}).$$

Using the delta method

$$T^{1/2}(\widehat{\gamma} - \gamma) \underset{T\to\infty}{\to} N\left(0_{m\times 1}\ ,\ V_\gamma\right)$$

where the covariance matrices $V_{\widetilde{\Phi}}$ and V_γ are linked by the expression

$$V_{\widetilde{\Phi}} = J_{\gamma,\phi} V_\gamma J_{\gamma,\phi}' \tag{52}$$

where

$$\begin{aligned} J_{\gamma,\phi} &= \frac{\partial g(\gamma)}{\partial \gamma'} = -D_{\gamma,\phi}^{-1} \times \frac{\partial f(\widetilde{\phi},\gamma)}{\partial \gamma'} \\ &= -D_{\gamma,\phi}^{-1} \times \frac{\partial f(\widetilde{\phi},\gamma)}{\partial vec(\Gamma)'} \times \frac{\partial vec(\Gamma)}{\partial \gamma'} = -D_{\gamma,\phi}^{-1} S_{\gamma,\phi} Q_\gamma \end{aligned} \tag{53}$$

and the functions $f(\widetilde{\phi},\gamma)$ and $g(\gamma)$ and the Jacobian $D_{\gamma,\phi} = D$ have been defined in the proof of Proposition 2 and $Q_\gamma = \partial q(\gamma)/\partial\gamma'$. The Jacobian $J_{\gamma,\phi}$ is $a \times m$ and has full column rank m. By solving (52) with respect to V_γ, yields the relation (32) in the text.■

References

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 (Chap. 3). Oxford: Blackwell.

Binder, M. and Pesaran, M. H. (1997), Multivariate linear rational expectations models. Characterization of the nature of the solutions and their fully recursive computation, *Econometric Theory* 13, 877-888.

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.

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

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.

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. (2007), Monetary policy with potentially misspecified models, *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.

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.

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

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

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

Hendry, D.F. (1976), The structure of simultaneous equations estimators, *Journal of Econometrics* 4, 51-88.

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

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.

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

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.

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.

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

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.

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.

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

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:									
ϖ_f	δ	γ_f	γ_b	ϱ	λ_y	λ_π	λ_r	ω_2	ω_3
0.25	0.10	0.30	0.70	0.13	0.5	1.5	0.5	-0.2	-0.3
Absolute value of largest eigen. of restricted VAR companion matrix $\widetilde{\Phi}^*$: 0.95									
Maximum likelihood estimates									
					T=200				
0.249	0.102	0.297	0.699	0.131	0.500	1.542	0.503	-0.205	-0.295
(0.045)	(0.028)	(0.042)	(0.042)	(0.034)	(0.088)	(0.151)	(0.047)	(0.046)	(0.031)
(0.127)*	(0.082)*	(0.089)*	(0.029)*	(0.054)*	(0.304)*	(0.440)*	(0.074)*	(0.042)*	(0.040)*
Frequency of rejections of LR test for CER: 0.094 (nominal level: 0.05)									
# of times in which an eigenvalue of $(\widehat{\Gamma}_0 - \widehat{\Gamma}_f\ \widehat{\widehat{\Phi}}_1)^{-1}\widehat{\Gamma}_f$ is found to be greater than one : 0									
					T=500				
0.247	0.100	0.298	0.699	0.131	0.500	1.516	0.502	-0.202	-0.299
(0.022)	(0.016)	(0.027)	(0.028)	(0.022)	(0.055)	(0.097)	(0.032)	(0.029)	(0.020)
(0.019)*	(0.051)*	(0.055)*	(0.031)*	(0.033)*	(0.188)*	(0.267)*	(0.046)*	(0.026)*	(0.024)*
Frequency of rejections of LR test for CER: 0.083 (nominal level: 0.05)									
# of times in which an eigenvalue of $(\widehat{\Gamma}_0 - \widehat{\Gamma}_f\ \widehat{\widehat{\Phi}}_1)^{-1}\widehat{\Gamma}_f$ is found to be greater than one : 0									

Table 1: Maximum likelihood estimates of the QR-DSGE model on simulated data. NOTES: Estimates are based on M=1000 simulated samples of length T, generated from the model (21), with k=3 and Gaussian disturbances. The value of parameters in the table are averages of 1000 FIML estimates; the numbers in parentheses without asterisks are the standard errors of the simulated distribution of 1000 estimates; the numbers in parentheses with asterisks are averages of the asymptotic standard errors obtained through the covariance matrix in (34). 100 samples have been discarded before starting computations, and each simulated sample is initiated with 100 additional observations to get a stochastic intial state, and these are then discarded. Zero values are used as starting values for structural parameters in the estimation. The LR test for CER is computed with reference to the 95 quantile from a χ^2 distribution with 27-10=17 degree of freedom, where 27 is the number of estimated coefficients of the unrestricted VAR (not including the covariance matrix), and 10 is the number of free parameters of the model.