Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN
Metodologia Statistica per la Ricerca Scientifica

XXVII Ciclo

Settore Concorsuale di afferenza: 13/A5

Settore Scientifico disciplinare: SECS-P/05

Estimation of Quasi-Rational DSGE Models

Presentata da
Giovanni Angelini

Coordinatore Dottorato
Prof. Angela Montanari

Relatore
Prof. Luca Fanelli

Esame finale anno 2015

Contents

List of Tables

List of Figures

1 Introduction

Small-scale dynamic stochastic general equilibrium models developed within the New Keynesian tradition (henceforth NK-DSGE models), have been treated as the benchmark of much of the monetary policy literature, given their ability to explain the impact of monetary policy on output, inflation and financial markets. 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 such as external habit persistence, implicit indexation, adjustment costs of investment, etc. (Christiano, Eichembaum and Evans, 2005; Smets and Wouters, 2003), these 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 still unclear and an increasing strand of the literature has become to put into question its usefulness in empirical modeling, see e.g. Pesaran and Smith (2011).

From the econometric point of view, DSGE models are interpreted as inherently misspecified systems (An and Schorfheide, 2007; Del Negro and Schorfheide, 2009; Canova and Ferroni, 2012) and are usually treated as restricted but parametrically incomplete representations of the data. The restrictions NK-DSGE place on VARs can be classified into two categories: (i) highly nonlinear cross-equation restrictions (CER) which the system places on its unique stable reduced form solution and which can be potentially used to recover estimates of the structural parameters; (ii) constraints on the lag VAR order which affect the state-space representation of the solution when some of the endogenous variables are not directly observable.

As regards (i) the problem is less apparent in a Bayesian framework. The recent frequentist literature suggests that the highly nonlinear nature of the CER may falsely led one to reject the structural model, when tested against the data, when asymptotic critical values are used in samples of lengths typically available to macroeconomists, especially when the variables are highly persistent, see e.g. Bekaert and Hodrick (2001), Cho and Moreno (2006), Li (2007), Fanelli (2008), Fanelli and Palomba (2011) and Bårdsen and Fanelli (2013). As regards (ii), Ravenna (2007), Fernández-Villaverde et al. (2007), Franchi and Vidotto (2013) and Franchi and Paruolo (2014) have provided the assumptions needed for a finite order VAR representation of a DSGE equilibrium model to exist. The coefficients of this VAR depend on the structural parameters through the CER (Komunjer and NG (2011), Iskrev (2010)). It usually happens that the constraints that the NK-DSGE model places on the VAR lag order are at odds with the dynamic features and persistence observed in quarterly (monthly) time series, inducing the omission of relevant propagation mechanisms that may characterize the data.

In this paper, we devote attention to the inherent omitted dynamics issue that characterizes the class of NK-DSGE models. We argue that the rational expectations paradigm may limit considerably the possibility of reproducing the actual autocorrelation structure of the data. Cole and Milani (2014) underline the [...failure of New Keynesian models under the rational expectations hypothesis to account for the dynamic interactions between observed macroeconomic expectations and macroeconomic realizations...]. Our idea is that in a world in which the data generating process is unknown and characterized by heterogeneous information sets, rational expectations are impossible to observe; multivariate time series models like VAR models can be regarded as 'boundedly rational' predictors 'in the spirit' of rational expectations, see Branch (2004). As is known, the 'standard' alternative to the rational expectations hypothesis in the literature is the adaptive learning hypothesis. Under adaptive learning, agents form their beliefs using forecast models with possibly time-varying coefficients and recursive updating rules through which they converge to an equilibrium can be achieved in the limit, see Evans and Honkapohja (2003) and references therein. Althought adaptive learning induces more persistence in the data (Branch and Evans 2006; Milani 2007; Chevillon *et al.* 2010) and permits a substantial statistical relaxation of the CER (Fanelli and Palomba 2011) compared to the rational expectations paradigm, the omitted dynamic issue is not solved within this paradigm.

We propose addressing the econometric analysis of NK-DSGE models under an alternative view which can be regarded as a an 'intermediate' position between rational expecations and learning, namely an adapted version of the quasi-rational expectations (QRE) hypothesis introduced by Nerlove *et al.*(1979), Nelson and Blessler (1992), Nerlove and Fornari (1998) and Holt and McKenzie (2003). The extreme form of QRE simply replaces expectational variables appearing in structural equations with their values calculated from the 'best fitting' reduced form model for them. However, since QRE are suited for exogenous, and not endogenous variables, we adapt the conventional concept to the specific framework of NK-DGSE models. The idea to apply the concept of Quasi-Rational to DSGE models is firstly proposed by Fanelli (2009).

We define the QR-NK-DSGE model as a linear rational expectations model derived from the baseline structural specification, such that its stable reduced form solution has the same lag structure as the state space (VAR) which fits the data optimally. This means that once a state space (VAR) is fitted to the observed time series, the QR-NK-DSGE model is obtained from the baseline structural specification by adding a number of auxiliary lags of the endogenous variables such that the implied unique stable solution, if it exists, corresponds to a restricted counterpart of the state space (VAR). The additional auxiliary parameters reflect the distance between actual agents' expectations and rational expectations. The number of lags characterizing the QR-NK-DSGE model is not arbitrary but depends explicity on the agent's forecast model. Other approaches, related to ours, have been suggested in the literature to improve the data adequacy of dynamic macro models based on RE. For instance, Kozicki and Tinsley (1999), Rudebusch (2002*a*, 2002*b*) and Fuhrer and Rudebusch (2004), introduce 'additional' dynamics and auxiliary parameters in the system they estimate, in recognition of the observed length of real world contracts, adjustment costs, delays in information flows and decision lags, see also Lindé (2005). In our setup, the additional auxiliary parameters estimated under QRE are motivated by the need of having a reduced form model solution consistent with the agents' statistical model. Compared to Curdia and Reis (2010), the advantage of using QR-NK-DSGE model in empirical analysis is that the zero restrictions in (ii) are automatically relaxed and therefore the risk of omitting relevant propagation mechanisms is under control. Clearly, there is no guarantee that the CER of type (i) are automatically fulfilled under QRE.

This paper is organized as follows. Section 1.1 introduces the Quasi Rational Expectations idea. In Section 1.2 is presented the reference model for our analysis. Section 2 discuss the case where all the endogenous variables are observed and Section 3 the case when one or more endogenous variables are unobserved. Section 5 presents two different applications based on the US Economy in the observeable and unobservable case. Some conclusions complete the paper.

1.1 Quasi Rational Expectations

Since the pioneristic work of Muth (1961), the main method to treat agents' expectations is the rational expectation (RE) hypothesis. Expectations have always played a central role in monetary policy because fluctuations in real activity and inflation are in large part driven by expectations about future demand, inflation and monetary and fiscal policy. Under this hypothesis, it is assumed that the agents know the Data Generating Process up to the unknown parameters. This approach has several limits, analyzed by Cole and Milani (2014) and Pesaran (1987): [...RE is not the way to modelling anything other than the steady state, because its informational assumption are too 'strong'...]. The princinpal reason for the failure of this approach is that the econometrician may fail to specify the behavioral model and the information available to the agents correctly. An alternative method which attempts to relax the 'basic' assumptions under RE is the Quasi Rational Expectations (QRE) hypothesis, discussed in Nerlove and Fornari (1998).[1] Following this approach, expectations are computed using the best fitting time series Autoregressive Integrated Moving Average ($ARIMA$) model for the variables. We focus on the following simple model

$$x_t = \gamma_f E_t x_{t+1} + \gamma_b x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0,1\right), \quad t = 1, ...T \qquad (1)$$

[1] The idea is originally proposed by Nerlove et al. (1979)

where the variable x_t is an observable scalar generated by a covariance stationary process, $E_t x_{t+1} = E(x_{t+1}|F_t)$ is the expectation operator conditional on the information set F_t and ε_t is a scalar white noise process with variance 1, called structural (or fundamental) disturbance, or also structural shock. We called the model in eq. (1) 'structural model'. The structural parameters are $\gamma_f > 0$,$\gamma_b > 0$ and are collected in $\theta = (\gamma_f, \gamma_b)'$. Assuming that $\gamma_f + \gamma_b < 1$ the unique stable RE solution of the eq. (1) is given by the Autoregressive process of order one ($AR(1)$)

$$x_t = \vartheta x_{t-1} + \psi \varepsilon_t, \quad t = 1, ...T \tag{2}$$

where $\vartheta = \vartheta(\theta)$ and $\psi = \psi(\theta)$ are reduced form parameters that depend nonlinearly on θ.These coefficients are subject to the cross equation restrictions

$$\gamma_f \vartheta^2 - \vartheta + \gamma_b = 0 \quad and \quad \psi = (1 - \gamma_f \vartheta)^{-1}$$

Under RE the data generating process belongs to the class of the mdels described by eq. (2).

Imagine now that the 'best fitting' time series model for x_t is given by an $AR(p)$ process with $p \geq 2$ and the DGP belongs to this class of models. Under this assumption, RE are not valid in the sense that the solution of model (1) is misspecified because it omits $p - 1$ relevant lags. Therefore RE solution omits some propagation mechanisms present in the DGP. This problem is anlyzed in detail in Section 2.1. One way to achieve such a result is suggested by Curdia and Reis (2010), they model the structure of ε_t using any time series process. Our idea, based on QRE, consists in rectifying the dynamic specification of the structural model such that its solution is consistent with the 'best fitting' time series model for the data. Consider for example the case where the 'best fitting' statistic model for x_t is an $AR(p=2)$

$$x_t = \vartheta_1 x_{t-1} + \vartheta_2 x_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0, \sigma_\varepsilon^2\right), \quad t = 1, ...T \tag{3}$$

where the autoregressive coefficient associated with the second lag, ϑ_2, is such that $\vartheta_2 \neq 0$. We call the model in eq. (3) the 'agents' statistical model (or forecast) model. The parameters of the statistical model are $\chi = (\vartheta_1, \vartheta_2, \sigma_\varepsilon^2)'$. Compared to the reduced form solution in eq. (2), the $AR(2)$ model in eq. (3) involves an additional lag of the variable x_t. For the econometrician, the best forecast of x_{t+1} conditional on the information set available at time t is $E\left(x_{t+1}|x_{t,}x_{t-1}, ..., x_1\right) = \vartheta_1 x_t + \vartheta_2 x_{t-1}$, not $E\left(x_{t+1}|x_{t,}x_{t-1}, ..., x_1\right) = \vartheta_1 x_t$.

Our goal is to find a reduced form solution whose number of lags is the same as in the 'best fitting' model. One way to impose that eq. (3) is the solution of the structural model is to redefine the model in eq. (1) by the following pseudo structural model

$$x_t = \gamma_f E_t x_{t+1} + \gamma_b x_{t-1} + \beta_2 x_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0, 1\right), \quad t = 1, ...T$$

where the parameter β_2 is an auxiliary parameter not derived from economic theory whose role is to fill up the gap between the theory and the dynamic of the data. Even if this parameter has no econimic meanig it is relevant in the estimation of the structural paramenters $\left(\gamma_f, \gamma_b\right)$. Indeed the solution of the new pseudo structural model is an $AR\,(2)$ in eq. (3) where the coefficients ϑ_1, ϑ_2 and σ_ε^2 are non linear function of the structural parameters and the auxiliary one (β_2) through the following CER

$$\gamma_f \left(\vartheta_1^2 + \vartheta_2\right) - \vartheta_1 + \gamma_b = 0 \tag{4}$$

$$\gamma_f \left(\vartheta_1 \vartheta_2\right) - \vartheta_2 + \beta_2 = 0 \tag{5}$$

$$\sigma_\varepsilon^2 = \left(1 - \gamma_f \vartheta_1\right)^{-2} \tag{6}$$

We can potentially use the CER in eq.s $(4) - (6)$ to recover the structural parameters. Compared to the model in eq. (2) in this case the dynamics of the system is consistent with the dynamic structure of the data.

1.2 Reference Model

We focus our analysis on the model proposed by Benati and Surico (2009). This model is used in several econometric analysis (Bardsen, and Fanelli,

2014) and it contains both forward and backward looking components:

$$\tilde{y}_t = \varpi_f E_t \tilde{y}_{t+1} + (1-\varpi_f)\tilde{y}_{t-1} - \delta(i_t - E_t\pi_{t+1}) + \upsilon_{y,t} \quad (7)$$

$$\pi_t = \frac{\beta}{1+\alpha\beta}E_t\pi_{t+1} + \frac{\alpha}{1+\alpha\beta}\pi_{t-1} + \varrho\tilde{y}_t + \upsilon_{\pi,t} \quad (8)$$

$$i_t = \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_t + \lambda_y \tilde{y}_t) + \upsilon_{i,t} \quad (9)$$

where

$$\upsilon_{x,t} = \rho_x \upsilon_{x,t-1} + \varepsilon_{x,t} \quad , -1 < \rho_x < 1 \ \ , \varepsilon_{x,t} \sim WN(0,\sigma_x^2) \ \ , \ \ x = \tilde{y}, \pi, i. \quad (10)$$

The variables $\tilde{y}_t = y_t - y_t^p$, π_t, and i_t stand for the output gap (y_t is output and y_t^p the natural rate of output), inflation, and the nominal interest rate, respectively; ϖ_f is the weight of the forward-looking component in the intertemporal IS curve; α is price setters' extent of indexation to past inflation; δ is households' intertemporal elasticity of substitution; β is a discount factor which is assumed to be fixed at the value $\beta = 0.99$; ϱ is the slope of the Phillips curve; λ_r, λ_π, and λ_y are the interest rate smoothing coefficient, the long-run coefficient on inflation, and that on the output gap in the monetary policy rule, respectively; finally, $\upsilon_{\tilde{y},t}$, $\upsilon_{\pi,t}$ and $\upsilon_{i,t}$ in eq. (10) are the mutually independent, autoregressive of order one and disturbances $\varepsilon_{\tilde{y},t}$, $\varepsilon_{\pi,t}$ and $\varepsilon_{i,t}$ are the structural (fundamental) shocks with variances σ_x^2, $x = \tilde{y}, \pi, R$. This or similar small-scale models have successfully been employed to conduct empirical analysis concerning the U.S. economy. Clarida *et al.* (2000)

and Lubik and Schorfheide (2004) have investigated the influence of systematic monetary policy over the U.S. macroeconomic dynamics; Boivin and Giannoni (2006), Benati and Surico (2009) have replicated the U.S. Great Moderation, while Castelnuovo and Fanelli (2014) have tested the determinacy/indeterminacy properties of the implied equilibria using identification-robust methods. Referring to the notation used in eq. (16), system (7) – (9) can be obtained by

$$\Gamma_0 = \begin{bmatrix} 1 & 0 & \delta \\ -\varrho & 1 & 0 \\ -(1-\lambda_r)\lambda_y & -(1-\lambda_r)\lambda_\pi & 1 \end{bmatrix}, \quad \Gamma_f = \begin{bmatrix} \varpi_f & \delta & 0 \\ 0 & \frac{\beta}{1+\alpha\beta} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{11}$$

$$\Gamma_b = \begin{bmatrix} (1-\varpi_f) & 0 & 0 \\ 0 & \frac{\alpha}{1+\alpha\beta} & 0 \\ 0 & 0 & \lambda_r \end{bmatrix} \tag{12}$$

We consider two different cases:

1. The vector $Z_t = [\tilde{y}_t, \pi_t, i_t]$ is assumed to be completely observed (Section 2). In this case the reduced form solution of this model is a VAR.

2. The variable y_t^p ($\tilde{y}_t = y_t - y_t^p$) is assumed to be unobserved (Section 3) and it follows a random-walk process $\left(y_t^p = y_{t-1}^p + \tau_{y,t}\right)$. In this case it is necessary a measurement equation to complete the system (7) – (9).

Under the random-walk assumption the measurement equation is the following (the derivation in Appendix A)

$$\Delta y_t = \Delta \tilde{y}_t + \tau_{y,t} \tag{13}$$

$$\pi_t = \pi_t \tag{14}$$

$$i_t = i_t \tag{15}$$

where Δ is the difference operator and $\tau_{y,t}$ is a measurement error. Differently from the previous case here the reduced form solution is a state-space model.

2 Observable variables: VAR representation

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 structural form of the NK-DSGE model is given

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \upsilon_t \tag{16}$$

where, $\Gamma_i = \Gamma_i(\theta)$, $i = 0, b, f$ are $n \times n$ whose elements depend on the vector of structural parameters θ, υ_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 $E_t Z_{t+1} = E(Z_{t+1} \mid \mathcal{F}_t)$.

The vector of structural parameters θ is a $n_\theta \times 1$ vector, the matrix Γ_0 is non-singular, while Γ_f and Γ_b can be singular and Γ_b possibly zero. When a direct link between the process generating υ_t and a set of 'forcing variables' is not provided by the theory, a typical completion of system (16) is obtained through the autoregressive specification

$$\upsilon_t = R\upsilon_{t-1} + \varepsilon_t \tag{17}$$

where R is a $n \times n$ diagonal stable matrix (i.e. with its eigenvalues - diagonal elements - inside the unit disk), and ε_t is a fundamental white noise term with covariance matrix Σ_ε. The assumption that the structural shocks are autocorrelated is common in the literature but is not generally derived from first-principles. Curdia and Reis (2010) model υ_t using an arbitrary time series model in order to capture the dynamics of the system.

There exists many solution methods available in the literature through which a reduced form solution of system $(16)-(17)$ can be computed under the rational expectations hypothesis, see among other Binder and Pesaran, (1995), Klein (2000) and Sims (2002). Remarkably, different solution methods can give rise to different representations. Assuming that a unique solution of the system exists, one way to express the reduced form solution associated with system $(16)-(17)$ is the VAR system

$$Z_t = \tilde{\Theta}_1 Z_{t-1} + \tilde{\Theta}_2 Z_{t-2} + \tilde{\Theta}_{dis}\varepsilon_t \tag{18}$$

where $\tilde{\Theta}_1 = \tilde{\Theta}_1(\theta)$, $\tilde{\Theta}_2 = \tilde{\Theta}_2(\theta)$ and $\tilde{\Theta}_{dis} = \tilde{\Theta}_{dis}(\theta)$ depend nonlinearly on θ through the CER (Appendix B)

$$(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - (\Gamma_b + R\Gamma_0) = 0_{n\times n} \tag{19}$$

$$(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_2 - \Gamma_f\tilde{\Theta}_1\tilde{\Theta}_2 + R\Gamma_b = 0_{n\times n} \tag{20}$$

$$(\Gamma_0 - \Gamma_f\tilde{\Theta}_1)^{-1} = \tilde{\Theta}_{dis}. \tag{21}$$

The matrix

$$\tilde{\Theta} = \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix}$$

which solves eq.s $(19)-(20)$ must be real and stable for the solution to be stable (asymptotically stationary). It can be proved (Castelnuovo and Fanelli, 2014) that the stability of the matrix $\Pi(\theta) = (\Gamma_0 + R\Gamma_f - \Gamma_f \tilde{\Theta}_1)^{-1}\Gamma_f$ is sufficient for uniqueness (determinacy).[2]

2.1 The ommitted dynamics issue

Given the NK-DSGE system (16) – (17), the reduced form model in eq. (18) and the CER in eq.s (19) – (21), the structural parameters θ can be estimated either in classical context through 'limited-' or 'full-information' methods, see e.g. Ruge-Murcia (2007) and Fukač and Pagan (2009) or via Bayesian methods, see among many other An and Schorfheide (2007).

With classical 'full-information' methods, one maximizes the likelihood function of the VAR system (18) subject to one of various approximations of the CER. For instance, one can maximize the likelihood function of the VAR in eq. (18) subject to the restrictions in eq.s (19) – (21) as in e.g. Cho and Moreno (2006) and Bårdsen and Fanelli (2014). These procedures, however, can fail to deliver consistent estimates of θ because of the misspecification of the VAR in eq. (18) with respect to the data, i.e. the possible omission of relevant lags.

Aside from the Bayesian solution suggested by Del Negro et al. (2007) and Del Negro and Schorfheide (2004) consisting in relaxing the strength

[2] The solution is not unique (i.e. there are multiple stable solutions) if $\tilde{\Theta}$ has eigenvalues inside the unit disk but the matrix $\Pi(\theta)$ has eigenvalues outside the unit disk, see Binder and Pesaran (1995), Section 2.3.

of the CER by using prior distributions for θ, classical approaches to cope with the poor dynamic structure implied by system (16) – (17) include the introduction of 'additional' dynamics in eq. (16) to account for real-word recognition, processing, adjustment costs and time-to-build lags as in Kozicki and Tinsley (1999), Rudebusch (2002a, 2002b) and Fuhrer and Rudebusch (2004) Lindé (2005, Section 5) and Jondeau and Le Bihan (2008), or the manipulation of the shock structure of υ_t in eq. (17) as in e.g. Smets and Wouters (2003, 2007) and Cúrdia and Reis (2010).

In this paper, with the term 'dynamics misspecification' we have in mind the situation in which the solution that captures the dynamics and persistence of Z_t satisfactorily involves more lags than the determinate reduced form solution associated with the NK-DSGE model in eq.s (18), see Consolo et al. (2009). A formalized definition is provided below.

To see how the postulated time series structure of the structural disturbance υ_t is related to the dynamics of the NK-DSGE model, notice that when in eq. (17) $R \neq 0_{n \times n}$, the reduced form in eq. (18) can be written as a (stable) constrained VAR of order two. Similarly, if υ_t in eq. (17) is specified as an autoregressive of order two, the implied determinate reduced form equilibrium of the NK-DSGE can be written as a constrained VAR of order three, and so forth.

Consider now an econometrician who observes $Z_1, Z_2, ..., Z_T$. Assume the econometrician finds that the 'best fitting' forecast model for Z_t is given by the VAR process

$$Z_t = A_1 Z_{t-1} + \cdots + A_k Z_{t-k} + A_{dis}\varepsilon_t \tag{22}$$

where $k \geq 3$, A_i, $i = 1, ..., k$ are $n \times n$ matrices of coefficients, ϵ_t is a white noise process with covariance matrix $\Sigma_\epsilon < \infty$ and A_{dis} is a $n \times n$ matrix. It can be recognized that under Assumption 1, the matching between the VAR coefficients in eq. (22) and the reduced form coefficients of the NK-DSGE model is given by

$$A_1 = \tilde{\Theta}_1, \tag{23}$$

$$A_2 = \tilde{\Theta}_2 \tag{24}$$

$$A_j = \tilde{\Theta}_j = 0_{n\times n}, \quad j = 3, 4, ..., k \tag{25}$$

$$A_{dis} = \tilde{\Theta}_{dis} \tag{26}$$

where $\tilde{\Theta}_1$and $\tilde{\Theta}_2$ are defined in eq.s (19) $-$ (21). There are two types of restrictions involved in eq.s (23) $-$ (26): (i) the constraints in eq.s (23) $-$ (24) and in eq. (26) which map the structural parameters in θ into the VAR coefficients (the $A_i s$ and A_{dis}); (ii) the zero restrictions in eq. (25) (and in eq. (24) when $R = 0_{n\times n}$) which reduce the VAR lag order from k to 2 (to 1 when $R = 0_{n\times n}$). Hence, there exists a discrepancy between the zero restrictions of the type (ii) and the idea that the VAR with lag order $k \geq 2$ fits the data optimally. This means the agents' best fitting model can not be regarded as the reduced form solution of the NK-DSGE model in eq.s $(16)-(17)$, unless the time series structure of v_t is adapted *ad hoc*. This observation leads to the following definition.

Definition 1 [Omitted dynamics] If the DGP belongs to the class of VAR models in eq. (22) , the NK-DSGE model summarized in eq.s $(16)-(17)$ entails an 'omitted dynamics issue' when the CER on the unique stable solution give rise to a set of zero-restrictions of the type in eq. (25) that reduces the lag order of the 'best-fitting' VAR for Z_t.

We define the QR-NK-DSGE model as a dynamic counterpart of the NK-DSGE specification in eq.s $(16)-(17)$ which circumvents the omitted dynamics issue of Definition 1. The solution of QR-NK-DSGE model has the same dynamics structure as the 'best fitting' model for Z_t. A formal definition will be given next.

2.2 The QR-NK-DSGE model

We have now all the ingredients to describe what we call QR-NK-DSGE model. Consider the VAR for Z_t

$$Z_t = A_1 Z_{t-1} + \cdots + A_k Z_{t-k} + \epsilon_t, \quad \epsilon_t \sim WN(0,\Sigma_\epsilon), \quad t = 1, ..., T \qquad (27)$$

where A_j, $j = 1, ..., k$ are $n \times n$ matrices of parameters and ϵ_t is a $n \times 1$ white noise process with covariance matrix Σ_ε; Z_0, Z_{-1}, ..., Z_{1-k} are fixed.

We consider the following assumptions:

Assumption 1 [Agents' forecast model] System (27) is the agents' forecast model and is such that $A_k \neq 0_{p \times p}$, $k \geq 3$.

Assumption 2 [Stationarity] The roots, s, of $\det[A(s)] = 0$ are such that $| s | > 1$, where $A(L) = I_p - \sum_{j=1}^{k} A_j L^j$ is the characteristic polynomial, and L is lag operator.[3]

Assumption 3 [Population parameters invariance] The parameters in $(A_1, \cdots, A_k, \Sigma_\varepsilon)$ do not vary over time.

Assumption 1 plays a crucial role, it maintains that any model restriction which reduces the VAR lag order generates the omitted dynamics issue of Definition 1. It is worth noting that in the jargon of the adaptive learning literature Assumption 1 might be re-stated by observing that the agents' Perceived Law of Motion is a VAR model for Z_t not necessarily restricted to be a Minimum State Variable solution, see e.g. Milani (2007). Assumption 2 rules out the occurrences of unit roots and explosive roots. However is possible to extend the analysis to the case of unit roots. Assumption 3 postulates time invariant parameters and guarantees that the mapping between the reduced form and the structural parameters is continuous. This hypothesis can be opportunely relaxed, provided the estimation of the structural

[3] In Appenx D a brief discussion if this assumption is not valid.

model is carried out over 'stable' sub-samples.[4] Assumption 3 is at odds with the adaptive learning hypothesis in which the (population) parameters of the perceived law of motion - the agents' beliefs - are generally treated as time-varying coefficients which are updated recursively as new information become available.

We propose the econometric analysis of small-scale DSGE models based on the QRE hypothesis. We look for a specification which reconciles the 'best' time series approximation of the data, given by the VAR in eq. (27), with the dynamic structure implied the DSGE model, without disregarding the nature of the CER that the latter imposes on the former.

Definition 2 [QR-DSGE Model] Under Assumptions 1-3 and the set of CER on the VAR in eq. (27), we define QR-NK-DSGE the pseudo-structural model defined by

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \sum_{j=2}^{k-1} \Phi_j Z_{t-j} \mathbb{I}_{\{k \geq 3\}} + \upsilon_t \quad (28)$$

$$\upsilon_t = R\upsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \Sigma_\varepsilon) \quad (29)$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function, Φ_j, $j = 2, ..., k-1$ is a $n \times n$ matrix containing additional parameters and ε_t is a white noise process with variance Σ_ε.

[4]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. Misspecification tests for structural instability play a crucial role in applied research. See also Juselius and Franchi (2007).

The matrices Φ_j can be diagonal or not, but the number of non-zero elements should not be too large to avoid the over-parameterization of the system. The number of matrices Φ_js is not arbitrary but is associated with the lag order of the best fitting VAR for the data in eq. (27). Let ϕ be the vector collecting the auxiliary parameters contained in the matrices Φ_j, $j = 2, ..., k-1$ and $\theta^* = (\theta', \phi')'$ the vector containing all (structural and auxiliary) parameters associated with the pseudo-structural form in eq.s $(28) - (29)$. Differently from Kozicki and Tinsley (1999), Rudebusch (2002a, 2002b), Fuhrer and Rudebusch (2004), the non-zero elements in Φ_j are not derived from the economic theory. We call the model pseudo-structural in the sense that the matrices Φ_js don't have a specific ecnomic meanig but their role is to fill the mismatch between the NK-DSGE equilibrium and the time series approximation of the data. Under Assumptions 1-3, if a unique reduced form solution for the model in eq. (28) exists, it is given by the VAR in eq. (27) with coefficients and disturbances subject to the restrictions $A_j = \tilde{\Theta}_j$, $j = 1, ..., k$, $\epsilon_t = \tilde{\Theta}_{dis}\varepsilon_t$, respectively, where the matrices $\tilde{\Theta}_j$ and $\tilde{\Theta}_{dis}$ are determined computing the eq. (16) under QRE, the CER are

(Appendix C)

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - \left(\Gamma_b + R\Gamma_0\right) = 0_{n\times n} \tag{30}$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_2 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_2 + \tilde{\Theta}_3\right) - \left(\Phi_2 - R\Gamma_b\right) = 0_{n\times n}$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_3 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_3 + \tilde{\Theta}_4\right) - \left(\Phi_3 - R\Phi_2\right) = 0_{n\times n}$$

$$\vdots$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_k - \Gamma_f\tilde{\Theta}_1\tilde{\Theta}_k + R\Phi_{k-1} = 0_{n\times n}$$

$$(\Gamma_0 - \Gamma_f\tilde{\Theta}_1)^{-1} = \tilde{\Theta}_{dis} \tag{31}$$

The solution is stable if the matrix $\breve{A}$ is stable (all the eigenvalues less than 1) where the matrix $\breve{A}$ is the restricted compact matrix of the solution

$$\breve{A} = \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 & \cdots & \tilde{\Theta}_k \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{n\times n} & 0_{n\times n} & I_n & 0_{n\times n} \end{bmatrix}.$$

3 Unobservable variables: state-space representation

In this section we extend the analysis of the QR-NK-DSGE model to the case where one or more endogenous variables of the system are unobserved. The solution of the model is the same as in eq. (18), a $VAR(2)$ for endogenous variables Z_t (the number of the VAR lags is determinated by eq. (17)). Differently from section 2, in this case the vector Z_t is not completely observed hence the $VAR(2)$ solution leads to a state space (or VARMA) model of the following form

$$\underbrace{\begin{bmatrix} Z_t \\ Z_{t-1} \end{bmatrix}}_{x_t} = \underbrace{\begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix}}_{A(\theta)} \underbrace{\begin{bmatrix} Z_{t-1} \\ Z_{t-2} \end{bmatrix}}_{x_{t-1}} + \underbrace{\begin{bmatrix} \tilde{\Theta}_{dis} \\ 0_{n\times n} \end{bmatrix}}_{S(\theta)} \varepsilon_t \tag{32}$$

$$y_t = Hx_t + N\tau_t \tag{33}$$

where $y_t = (y_{1,t}, y_{2,t}, \cdots, y_{m,t})'$ is the $m \times 1$ vector of observable variables and τ_t is a $n_\tau \times 1$ vector of the measurement errors. Eq. (32) is the state equation and eq. (33) is the measurement equation. The matrices $\tilde{\Theta}_1 = \tilde{\Theta}_1(\theta)$, $\tilde{\Theta}_2 = \tilde{\Theta}_2(\theta)$ and $\tilde{\Theta}_{dis} = \tilde{\Theta}_{dis}(\theta)$ depend nonlinearly on θ through the

CER (Appendix B)

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - \left(\Gamma_b + R\Gamma_0\right) = 0_{n\times n} \tag{34}$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_2 - \Gamma_f\tilde{\Theta}_1\tilde{\Theta}_2 + R\Gamma_b = 0_{n\times n}$$

$$(\Gamma_0 - \Gamma_f\tilde{\Theta}_1)^{-1} = \tilde{\Theta}_{dis}. \tag{35}$$

By substituting eq. (32) into eq. (33) and using some algebra one obtains the representation

$$\begin{aligned} x_t &= A(\theta)x_{t-1} + S(\theta)\varepsilon_t \\ y_t &= H\left(A(\theta)x_{t-1} + S(\theta)\varepsilon_t\right) + N\tau_t \\ &= HA(\theta)x_{t-1} + HS(\theta)\varepsilon_t + N\tau_t \end{aligned}$$

Now defining $\xi_t = \begin{bmatrix} \varepsilon_t & \tau_t \end{bmatrix}'$ and rearraging terms the system above becomes

$$x_t = A(\theta)x_{t-1} + B(\theta)\xi_t \tag{36}$$

$$y_t = C(\theta)\,x_{t-1} + D(\theta)\,\xi_t \tag{37}$$

where

$$\underset{2n\times(n+n_\tau)}{B(\theta)} = \begin{bmatrix} S(\theta) & 0_{2n\times n_\tau} \end{bmatrix}, \underset{m\times 2n}{C(\theta)} = HA(\theta), \underset{m\times(n+n_\tau)}{D(\theta)} = \begin{bmatrix} HS(\theta) & N \end{bmatrix}$$

The System (36) – (37) defines the so-called 'A, B, C (and D's)' representation of our NK-DSGE model, see Fernandez-Villaverde et al. (2007), Ravenna (2007) and Franchi and Paruolo (2012). In this case the identification issue is an important challenge, see among other Komijer and Ng (2011) and Guerron-Quintana et al (2013). The authors developed a set of rank condions using the matrices A,B,C and D in the eq.s (36) – (37).

3.1 The Kalman filter approach

The Kalman Filter (henceforth KF) is a useful tool, proposed by Kalman (1960), for evaluating the likelihood function in State Space models. Consider a standard State Space model (as in eq. (32) – (33)):

$$State\ Equation : x_t = Ax_{t-1} + S\varepsilon_t \tag{38}$$

$$Measurement\ Equation : y_t = Hx_t + N\tau_t \tag{39}$$

where x_t is a $n \times 1$ vector of endogenous variables and y_t is a $m \times 1$ vector of the observable. In presence of unobservable components a possible way to compute the likelihood of the system in eq.s (38) – (39) is the KF. Starting with initial values for $x_0 \sim N\left(x_{0|0}, P_{0|0}\right)$ and assuming normality for the error terms:

$$\varepsilon_t \sim N(0,Q) \tag{40}$$

$$\tau_t \sim N(0,R) \tag{41}$$

is possbile to compute the likelihood function. The KF is a recursive algorithm composed by two steps: forecasting and updating. The forecasting step starts with the initial value $x_{0|0}$ and using eq. (38) derives the values for $x_{1|0} = Ax_{0|0}$ and $Var\left(x_{1|0} \mid x_{0|0}\right) = P_{1|0} = AP_{0|0}A' + SQS'$. Using eq. (39) is possible to obtain the values $y_{1|0} = Hx_{1|0} + N$ and $V_{1|0} = HP_{1|0}H' + NRN'$, where $V_{t|t-1} = Var\left(y_t \mid y_1...y_{t-1}, \theta\right)$. At this point, the updating step starts. Normality in eq.s $(40) - (41)$ leads to a Gaussian process so the distribution of x and y is given by:

$$\begin{bmatrix} y_1 \\ x_1 \end{bmatrix} \sim N\left(\begin{bmatrix} y_{1|0} \\ x_{1|0} \end{bmatrix}, \begin{bmatrix} V_{1|0} & HP_{1|0} \\ P_{1|0}H' & P_{1|0} \end{bmatrix}\right)$$

hence, from the property of the multivariate normal distribution x and P are equal to

$$x_{1|1} = x_{1|0} + P'_{1|0}H'V_{1|0}^{-1}\left(y_1 - y_{1|0}\right) \tag{42}$$

$$P_{1|1} = P_{1|0} - P'_{1|0}H'V_{1|0}^{-1}HP_{1|0} \tag{43}$$

It is now necessary to repeat the forecasting and updating step for $t = 2...T$ and the likelihood function is given by

$$L\left(y \mid \theta\right) = \prod_{t=1}^{T} \left(2\pi\right)^{-\frac{m}{2}} \left|V_{t|t-1}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(y_t - y_{t|t-1}\right)' V_{t|t-1}^{-1}\left(y_t - y_{t|t-1}\right)\right\} \tag{44}$$

3.1.1 Kalman Smoother

If the kalman filter is a tool to compute the likelihood function in presence of unobservable components, Kalman Smoother (heceforth KS) is a tool to correct the values of the unobservable variables x_t after the computation of the KF. Starting from the last values obtained in the KF procedure $x_{t|T}$ and $P_{t|T}$, KS corrects all the values for the unobservable variables. The algorithm starts from the observation $T-1$ (obviously the last value will not be corrected because it contains already the whole information and this values are $x_{T|T}$, $P_{T|T}$ obtained in the eq.s $(42)-(43)$) and proceedes backwards until the first observation, so for each $t = (T-1), \ldots, 1$ the value for the state and for the variance is given by

$$x_{t|T} = x_{t|t} + P_{t|t}' A' P_{t+1|t}^{-1}\left(x_{t+1|T} - x_{t+1|t}\right) \tag{45}$$

$$P_{t|T} = P_{t|t} - P_{t|t}' A' P_{t+1|t}^{-1}\left(P_{t+1|t} - P_{t+1|T}\right) P_{t+1|t}^{-1} A P_{t|t}' \tag{46}$$

where $x_{t+1|t} = Ax_{t|t}$ and $P_{t+1|t} = AP_{t|t}A' + SQS'$.

3.2 The omitted dynamic issue

The state space system in eq.s (36) – (37) summarizes the determinate equilibrium associated with the NK-DSGE model under RE, and collapses to a VAR for Z_t when $y_t = Z_t$. Provided θ is locally identifiable, the state space model in eq.s (36) – (37) can be taken to the data using different estimation methods, see e.g. Ruge-Murcia (2007), DeJong and Dave (2007), Fukač and Pagan (2009) and Bårdsen and Fanelli (2014). These procedures, however, can fail to deliver consistent estimates of θ when the omitted dynamics issue occurs, see e.g. Jondeau and Le Bihan (2008).

To characterize the omitted dynamic issue, assume that the agents' 'best fitting' model for the data can be represented by the state space system

$$\underbrace{\begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-k+1} \end{bmatrix}}_{x_t^*} = \underbrace{\begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \ddots & \vdots \\ 0_{n\times n} & \cdots & I_n & 0_{n\times n} \end{bmatrix}}_{A^*} \underbrace{\begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-k} \end{bmatrix}}_{x_{t-1}^*} + \underbrace{\begin{bmatrix} A_{dis} \\ 0_{n\times n} \\ 0_{n\times n} \\ 0_{n\times n} \end{bmatrix}}_{S^*} \epsilon_t \quad (47)$$

$$y_t = Hx_t^* + N\tau_t \quad (48)$$

where A_j, $j = 1, ..., k$ are $n \times n$ matrices of parameters ϵ_t is a $n \times 1$ white noise process with covariance matrix Σ_ϵ. By setting $\epsilon_t = A_{dis}\varepsilon_t$, with A_{dis} a $n \times n$ matrix, it can be recognized that under Assumption 1, the matching between the state-space coefficients in eq.s (47) – (48) and the reduced form coefficients of the NK-DSGE in eq.s $(32)-(33)$ is given by

$$\begin{aligned} A_1 &= \tilde{\Theta}_1, \\ A_2 &= \tilde{\Theta}_2 \\ A_j &= \tilde{\Theta}_j = 0_{n\times n}, \quad j=3,4,...,k \\ A_{dis} &= \tilde{\Theta}_{dis} \end{aligned} \tag{49}$$

Where $\tilde{\Theta}_1$, $\tilde{\Theta}_2$ and $\tilde{\Theta}_{dis}$ are subjected to the CER in eq.s $(19)-(20)$.We define the QR-NK-DSGE model as a dynamic counterpart of the NK-DSGE specification which circumvents the omitted dynamics issue of Definition 1 in eq. (49).

3.3 The QR-NK-DSGE model

Consider now that the 'best fitting' state-space model for the data is given in eq.s $(47)-(48)$ and the Assumptions 1-3 are still valid the Quasi-Rational 'pseudo-structural form' is given by

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \sum_{j=2}^{k-1} \Phi_j Z_{t-j} \mathbb{I}_{\{k\geq 3\}} + \upsilon_t \tag{50}$$

$$\upsilon_t = R\upsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0, \Sigma_\varepsilon\right) \tag{51}$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function, Φ_j, $j=2,...,k-1$ is a $n\times n$ matrix containing additional parameters and ε_t is a white noise process with variance

Σ_ε. When $k \leq 2$ (k is the number of lags in the best fitting statistical model in eq.s $(47)-(48)$) the pseudo structural form collapses to the NK-DSGE model under RE. To fully understand the nature of the system in eq.s $(50)-(51)$, observe that $k \geq 3$, its i-th Euler equation reads

$$\begin{aligned} Z_{i,t} &= \gamma_{i,0}' Z_{i,t}^* + \gamma_{i,f}' E_t Z_{t+1} + \gamma_{i,b}' Z_{t-1} + \left(\sum_{j=2}^{k-1} \phi_{i,j} Z_{i,t-j} \right) + \upsilon_{i,t} \\ \upsilon_{i,t} &= R_i \upsilon_{i,t-1} + \varepsilon_{i,t}, \quad i = 1, ..., n. \end{aligned}$$

In this equation, the $(n-1) \times 1$ vector $Z_{i,t}^*$ denotes Z_t with its i-th entry suppressed, the $(n-1) \times 1$ vector $\gamma_{i,0}$ collects the structural parameters that enter the i-th row of Γ_0, the $n \times 1$ vector $\gamma_{i,f}$ collects the structural parameters that enter the i-th row of Γ_f, the $n \times 1$ vector $\gamma_{i,b}$ contains the structural parameters that enter the i-th row of Γ_b, $\phi_{i,j}$ is the i-th diagonal element of Φ_j, $j = 1, ..., k-1$ and, finally, $\upsilon_{i,t}$ and $\varepsilon_{i,t}$ are the i-th elements of the vectors υ_t and ε_t, respectively, where the autoregressive parameter -1 $< R_i <$ 1 is the i-th diagonal component of R.

Let ϕ be the vector collecting the auxiliary parameters contained in the matrices Φ, $j = 2, ..., k-1$ and $\theta^* = (\theta', \phi')'$ the vector containing all (structural and auxiliary) parameters associated with the pseudo-structural form in eq.s $(50)-(51)$. Under Assumptions 1-3, if a unique reduced form solution for the model in eq.s $(50)-(51)$ exists, it is given by the state-space model in eq.s $(47)-(48)$ with coefficients and disturbances subject to the restrictions

$A_j = \tilde{\Theta}_j$, $j = 1, ..., k$, $A_{dis} = \tilde{\Theta}_{dis}$ and $\tilde{\Theta}_{j,} j = 1, ..., k$ and $\tilde{\Theta}_{dis}$ fullfil the CER (Appendix C)

$$\begin{aligned}
(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - (\Gamma_b + R\Gamma_0) &= 0_{n\times n} \\
(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_2 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_2 + \tilde{\Theta}_3\right) - (\Phi_2 - R\Gamma_b) &= 0_{n\times n} \\
(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_3 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_3 + \tilde{\Theta}_4\right) - (\Phi_3 - R\Phi_2) &= 0_{n\times n} \\
&\vdots \\
(\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_k - \Gamma_f\tilde{\Theta}_1\tilde{\Theta}_k + R\Phi_{k-1} &= 0_{n\times n}
\end{aligned} \tag{52}$$

$$(\Gamma_0 - \Gamma_f\tilde{\Theta}_1)^{-1} = \tilde{\Theta}_{dis} \tag{53}$$

the solution is stable if the matrix $\breve{A}$ is stable (all the eigenvalues less than 1), where the matrix $\breve{A}$ is the restricted compact matrix of the solution

$$\breve{A} = \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 & \cdots & \tilde{\Theta}_k \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \vdots & \vdots \\ 0_{n\times n} & 0_{n\times n} & I_n & 0_{n\times n} \end{bmatrix}$$

4 Estimation procedure

The econometrics analysis of the NK-DSGE model based on QRE can be based on the following two steps:

Step 1 Fit the statistical model to the data and use available information criteria or likelihood-ratio tests to determine the lag lenght k that fits the data optimally. The statistical model is eq. (27) if Z_t is completely observable or eq.s (47) $-$ (48) if one or more component in Z_t are not observed. This can be done by estimating the VAR or the state space model with the maximum likelihood approach. In the observable case, given the following 'best fitting model'

$$Z_t = \bar{A}\bar{Z}_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0,\Sigma_\epsilon)\ ,\ t = 1, ..., T$$

where $\bar{A} = [A_{1,}..., A_k]$ and $\bar{Z}_{t-1} = [\bar{Z}_{t-1}, ..., \bar{Z}_{t-k}]'$, the log-likelihood (given ν the vector of parameters) is the following

$$\ln \mathcal{L}(Z \mid \nu) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log(|\Sigma_\epsilon|) - \frac{1}{2}\sum_{t=1}^{T}\left(\left(Z_t - \bar{A}\bar{Z}_{t-1}\right)'\Sigma_\epsilon^{-1}\left(Z_t - \bar{A}\bar{Z}_{t-1}\right)\right) \quad (54)$$

In the unobservable case the 'best statistical model' is given

$$\underbrace{\begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-k+1} \end{bmatrix}}_{x_t^*} = \underbrace{\begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \ddots & \vdots \\ 0_{n\times n} & \cdots & I_n & 0_{n\times n} \end{bmatrix}}_{A^*} \underbrace{\begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-k} \end{bmatrix}}_{x_{t-1}^*} + \underbrace{\begin{bmatrix} A_{dis} \\ 0_{n\times n} \\ 0_{n\times n} \\ 0_{n\times n} \end{bmatrix}}_{S^*} \epsilon_t$$

$$y_t = Hx_t^* + N\tau_t$$

and the log-likelihood is computed by the Kalman Filter procedure in Section (3.1) and is

$$\ln \mathcal{L}(Z \mid \nu) = -\frac{mT}{2}\log(2\pi) + \sum_{t=1}^{T}\left(-\frac{1}{2}\log\left(\left|V_{t|t-1}\right|\right) - \frac{1}{2}\left(y_t - y_{t|t-1}\right)' V_{t|t-1}^{-1}\left(y_t - y_{t|t-1}\right)\right) \tag{55}$$

where $V_{t|t-1}^{-1}$ and $y_{t|t-1}$ are computed during the Kalman Filter algorithm and ν is the vector of parameters. In this case the specific procedure we use to find a global maximum is the simulated annealing/genetic algorithm of Andreasen (2010). For each estimated model, we check whether the minimality (controllability and observability) and local identification conditions discussed in Komunjer and Ng (2011) are satisfied in correspondence of the maximum likelihood estimation.

Step 2 Given k, estimate $\theta^* = (\theta', \phi')'$ from the statistical model (eq. (27) in the observable case and eq.s $(47)-(48)$ in the unobservable case) under a numerical approximation of the CER in eq.s $(30) - (31)$ or in eq.s $(52) - (53)$. This can be done by using a bayesian algorithm sketched in Section 4.1 or the algorithm of Andreasen (2010) described in Section 4.2.

4.1 Bayesian approach

Given a set of priors for the parameters in θ^* is given $(p(\theta^*))$, the posterior distribution is computed using an Markov Chain Monte Carlo (MCMC) algorithm as the Random-Walk Metropolis (RWM). The sketch of this algorithm is the following (see Del Negro and Schorfheide (2012) or An and Schorfheide (2007) for more details).

Algorithm 1 *RWM*

1. *Use a numerical optimization routine to maximize the log posterior, which up to constant is given by* $\ln \mathcal{L}(Z \mid \theta^*) + \ln p(\theta^*)$*, where the term* $\mathcal{L}(Z \mid \theta^*)$ *indicate del log-likelihood and* $p(\theta^*)$ *is the joint prior. Denote the posterior mode by* $\tilde{\theta}$.

2. *Let* $\tilde{\Sigma}$ *the inverse of the (negative) Hessian computed at the posterior mode* $\tilde{\theta}$*, which can be computed numerically.*

3. *Draw* $\theta^{*(0)}$ *from* $N\left(\tilde{\theta}, \tilde{\Sigma}\right)$ *or directly specify a starting value.*

4. *For* $j = 1, ..., W$*: draw* ϑ *from the proposal distribution* $N\left(\theta^{*(j-1)}, c^2\tilde{\Sigma}\right)$. *The new candidate parameter vector* ϑ *is accepted* $\left(\theta^{*(j)} = \vartheta\right)$ *with*

probability min $\{1, r\left(\theta^{*(j)}, \vartheta \mid Y\right)\}$ *and rejected* $\left(\theta^{*(j)} = \theta^{*(j-1)}\right)$ *otherwise. Here:* $r\left(\theta^{*(j)}, \vartheta \mid Y\right) = \frac{\mathcal{L}(Y|\theta^{*})p(\theta^{*})}{\mathcal{L}(Y|\theta^{*(j-1)})p(\theta^{*(j-1)})}$*.The choiche of the costant c is an important issue in bayesian estimation, usually this parameter is chosen to reach an acceptance ratio close to 20% (acceptance ratio is given dividing the number of accepted candidates for the total number of iterations).*

5. *The posterior distribution of* θ^{*} *is given by the set of accepted* $\theta^{*(j)}, j = 1..., W$.

To compare the models in a Bayesian approach we use two standard criteria commonly used in the Bayesian model comparison framework: Deviance Information Criterion (DIC) and Predictive Model Choice Criterion (PMCC). DIC is a generalization of the AIC and BIC criteria used in MCMC algorithms. This indicator is defined by

$$DIC = 2 \cdot \bar{D} - D\left(\bar{\theta}\right)$$

where $\bar{\theta}$ is the expected value of θ^{*}, $D\left(\bar{\theta}\right) = -2\log\left(\mathcal{L}(Z \mid \bar{\theta})\right)$ and $\bar{D} = E\left[D\left(\theta^{*}\right)\right]$. The lowest DIC is associated with the best model.

PMCC proposed by Gelfand and Gosh (1998) is defined

$$PMCC = \sum_{i=1}^{n}\left(\mu_i - Z_i\right)^2 + \sum_{i=1}^{n}\sigma_i^2$$

where $\mu_i = E(Z_{rep,i} \mid Z)$, $\sigma_i^2 = Var(Z_{rep,i} \mid Z)$ and Z_{rep} is the posterior predictive distribution. Similarly to the DIC, the model with lower PMCC is preferred.

4.2 Frequentist approach

A possible method which can be used to find the global maximum of the likelihood for DSGE models is the algorithm developed by Hansen et al. (2005) and recently applied to the context of DSGE models by Andreasen (2010), called CMA-ES. Andreasen shows that with ten structural parameters, this routine finds the global optimum in the 95% of cases (different initial values) and with 20 and 35 parameters the routine finds the global optimum in 85% and 71% of cases respectively. Before applying this procedure to the data, we run a Monte Carlo simulation study to envisage how the Andreasen's algorithm works in our setup. Two different simulations are proposed: one for the RE case and one for the QRE case. Table 1 reports the results of the simulation in the RE case. The model is proposed in Section 1.2, considering Z_t not completely completely observed and the 'true' values of the parameters (reported in the first column of Table 4.2) are taken by Benati and Surico (2009)

Table 1. Monte Carlo (RE)

True parameters	$mean(std.err)$
$\varpi_f = 0.744$	0.736(0.107)
$\delta = 0.124$	0.132(0.061)
$\alpha = 0.059$	0.058(0.046)
$\varrho = 0.044$	0.047(0.035)
$\lambda_r = 0.834$	0.822(0.067)
$\lambda_\pi = 1.749$	1.956(1.001)
$\lambda_y = 1.146$	1.355(0.587)
$\rho_y = 0.796$	0.776(0.137)
$\rho_\pi = 0.418$	0.406(0.117)
$\rho_i = 0.404$	0.407(0.157)
$\sigma_y^2 = 0.055$	0.066(0.054)
$\sigma_\pi^2 = 0.391$	0.040(0.138)
$\sigma_i^2 = 0.492$	0.523(0.155)
$\sigma_{\tau_y}^2 = 0.250$	0.271(0.273)
$LR\ (\alpha = 0.05), rejection\ rate = 6.3\%$	
$M = 150, T = 100$	

Notes: M is the number of simulations and T is the sample size. In brachets the Monte Carlo standard errors. The LR test is computed between the DSGE and the unrestricted State Space.

In the first column of Table 4.2 there are the true parameters and in the second column the mean over the $M = 150$ simulations. The data for the Monte Carlo simulation study are generated under the 'true model' given by the eq.s $(32) - (33)$ where $\tilde{\Theta}_1, \tilde{\Theta}_2$ and $\tilde{\Theta}_{dis}$ are computed using Binder and Pesaran (1995) method. For each simulation (different data) we estimate the model in eq.s $(32) - (33)$ subject to the CER in eq.s $(34) - (35)$ and the unrestricted model in eq.s $(47) - (48)$ with $k = 2$. For each case then we compare, through a LR test (with $\alpha = 0.05$), this two models and the rejection rate is the percentage of the whole simulation study for which we reject the null model (the model in eq.s (32)-(33) subject to the CER is nested to the unrestricted model in eq.s $(47) - (48)$). The likelihood ratio test (LR) is given by

$$LR = -2 \left(\log \left(Likelihood^{null} \right) - \log \left(Likelihood^{alternative} \right) \right)$$

and is asymptotically distribuited as a $\chi^2_{(q_{alternative} - q_{null})}$ where $q_{alternative}$ and q_{null} are the free parameters in the alternative model and null model.

Table 4.2 reports the results of the Monte Carlo simulation study in the QRE case. The 'true' values of the parameters in the first column are selected in such a way that an unique and stable solution exist.

Table 2. Monte Carlo (QRE)

True parameters	$mean(std.err)$
$\varpi_f = 0.744$	0.593(0.211)
$\delta = 0.124$	0.126(0.046)
$\alpha = 0.059$	0.061(0.029)
$\varrho = 0.044$	0.055(0.028)
$\lambda_r = 0.834$	0.834(0.085)
$\lambda_\pi = 2.200$	2.918(1.649)
$\lambda_y = 1.150$	1.096(0.592)
$\rho_y = 0.796$	0.617(0.251)
$\rho_\pi = 0.418$	0.394(0.107)
$\rho_i = 0.404$	0.376(0.143)
$\phi_{1,2} = 0.060$	-0.009(0.127)
$\phi_{2,2} = -0.500$	-0.509(0.169)
$\phi_{3,2} = 0.100$	0.070(0.124)
$\sigma_y^2 = 0.055$	0.043(0.034)
$\sigma_\pi^2 = 0.391$	0.368(0.058)
$\sigma_i^2 = 0.492$	0.478(0.079)
$\sigma_{\tau_y}^2 = 0.250$	0.061(0.044)
$LR\ (\alpha = 5\%), rejection\ rate = 6.4\%$	
$M = 50, T = 100$	

Notes: M is the number of simulations and T is the sample size. In bracket the Monte Carlo standard errors. The LR test is computed between the DSGE and the unrestricted State Space.

In both simulations the results are quite good: the mean among the simulations is very close to the true values for each parameters and the rejection rate is similar to the nominal value $(\alpha = 5\%)$. The only problematic parameter is $\phi_{1,2}$. The difference between the true value (0.060) and the estimated one (-0.009) is probably due to the small sample size.

In addition to the LK test for the model comparison in the frequentist case we also use the standard information criteria.

$$
\begin{aligned}
Akaike &= -2\log\left(Likelihood\right) + 2q \\
Hannan - Quinn &= -2\log\left(Likelihood\right) + 2q\log\left(\log\left(T\right)\right) \\
Schwarz &= -2\log\left(Likelihood\right) + q\log\left(T\right)
\end{aligned}
$$

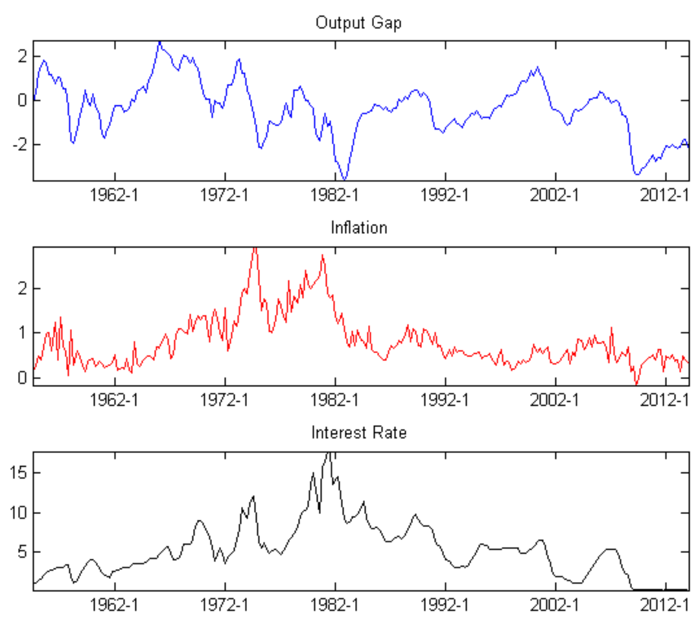
where q is the number of parameters and T is the sample size.

5 Empirical applications

5.1 Observable case

In this section we estimate a QR-NK-DSGE model for the US economy assuming that the vector $Z_t = [\tilde{y}_t, \pi_t, i_t]'$ is completely observed. The output gap series $\tilde{y}_t$ is given by the difference between real GDP (y_t) and potential output proxied by Congressional Budget Office (CBO) "output gap" series. The inflation rate π_t is measured by the log of the quarterly changes in the GDP deflator and i_t is the short-term nominal interest rate.

Figure1. Time Series



As common for quarterly data, the discount factor parameter β is fixed to 0.99 (corresponding to an annual discount rate of approximately 4%). We focus on the model discussed in Section 1.2 and drop out the period before the Volcker stabilization (1979) and after the recent financial crysis (2008:Q4) from the analysis. In this period, the macroeconomic literature has documented a dramatic fall in the variances of the main macroeconomic indicators, which has been termed the 'Great Moderation'. McConnell and Perez-Quiros (2000) identify 1984:Q1 as the break-date of the variance of the U.S. real GDP, so we consider the period 1984:Q2-2008:Q3 ($T = 98$). According to the analyses in Castelnuovo and Fanelli (2014), it is possible to reject the occurence of sunspot-driven equilibria and not to reject the occurence of a model's solution consistent with the unique equilibrium in that period. The 'best-fitting' statistical model for Z_t is a VAR with $k = 3$ lags

$$Z_t = A_1 Z_{t-1} + A_2 Z_{t-2} + A_3 Z_{t-3} + \epsilon_t, \quad \epsilon_t \sim WNN\left(0, \Sigma_\epsilon\right) \tag{56}$$

Table 3 summerizes the criteria for the lag lenght selection. The likelihood for the unrestricted model is the one in eq. (54) and it is computed for the lags $k = 1, ..., 6$. For the lag length selection we consider the likelihood ratio test and the standard information criteria (we select $k_{\max} = 6$ because 6 lags are enough to capture the dynamics in the data)

Table 3. Lag lenght selection

lag	Likelihood	LR	p-value	Akaike	Hannan-Quinn	Schwartz
1	27.08	158.45	< 0.001	-30.16	-17.67	0.73
2	55.23	102.14	< 0.001	-68.47	-46.70	-14.61
3	86.82	38.98	0.063^*	-113.63^*	-82.67^*	-37.01^*
4	94.34	23.93	0.157	-110.67	-70.61	-11.49
5	100.32	11.98	0.214	-104.65	-55.55	16.93
6	106.31	–	–	-98.61	-40.59	45.13

Notes: the log-likelihood is maximized using the standard economitrician analysis for VAR models. The LR tests are computed by comparing the log-likelihoods obtained with k=2,...,6 lags and the log-likelihood obtained with k=6 lags. Asterisks denote the optimallag selected by the test/criterion. The sample is 1984:Q2-2008:Q3 and the model is described in eq. (56).

Given the results in Table 3, we select $k = 3$ as the optimal lag. Indeed each information criterion and the LR test select $k = 3$. From Definition 2, the QR-NK-DSGE model is given by the system

$$\begin{aligned} \Gamma_0 Z_t &= \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \Phi_2 Z_{t-2} + \upsilon_t \\ \upsilon_t &= R\upsilon_{t-1} + \varepsilon_t, \varepsilon_t \sim WN\left(0, \Sigma_\varepsilon\right) \end{aligned} \tag{57}$$

where

$$\Phi_2 = \begin{bmatrix} \phi_{1,2} & 0 & 0 \\ 0 & \phi_{2,2} & 0 \\ 0 & 0 & \phi_{3,2} \end{bmatrix}, R = \begin{bmatrix} \rho_y & 0 & 0 \\ 0 & \rho_\pi & 0 \\ 0 & 0 & \rho_i \end{bmatrix}.$$

The CER in this case are summarized in eq. $(30) - (31)$. The vector of structural parameters, including the auxiliary ones, is $\theta^* = (\theta, \phi')'$, where $\theta = (\delta, \varpi_f, \alpha, \beta, \varrho, \lambda_r, \lambda_\pi, \lambda_y)', \phi = (\phi_{1,2}, \phi_{2,2}, \phi_{3,2})'$. The pseudo-structural equations are

$$\tilde{y}_t = \varpi_f E_t \tilde{y}_{t+1} + (1-\varpi_f)\tilde{y}_{t-1} - \delta(i_t - E_t\pi_{t+1}) + \phi_{1,2}\tilde{y}_{t-2} + \upsilon_{y,t} \quad (58)$$

$$\pi_t = \frac{\beta}{1+\alpha\beta} E_t\pi_{t+1} + \frac{\alpha}{1+\alpha\beta}\pi_{t-1} + \varrho\tilde{y}_t + \phi_{2,2}\pi_{t-2} + \upsilon_{\pi,t} \quad (59)$$

$$i_t = \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_t + \lambda_y \tilde{y}_t) + \phi_{3,2} i_{t-2} + \upsilon_{i,t} \quad (60)$$

$$\upsilon_{x,t} = R\upsilon_{x,t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0, \Sigma_\varepsilon\right), \quad x = \tilde{y}, \pi, i. \quad (61)$$

The Step 2 of the estimation procedure summarized in Section 4 requires estimating $\theta^* = (\theta, \phi')'$ from the model in eq. (56) by imposing the CER $A_j = \tilde{\Theta}_j$, $j = 1,2,3$, $\Sigma_\epsilon = \tilde{\Sigma}_\epsilon$, where

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - \left(\Gamma_b + R\Gamma_0\right) = 0_{3\times 3} \quad (62)$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_2 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_2 + \tilde{\Theta}_3\right) - \left(\Phi_2 - R\Gamma_b\right) = 0_{3\times 3}$$

$$\left(\Gamma_0 + R\Gamma_f\right)\tilde{\Theta}_3 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_3\right) + R\Phi_2 = 0_{3\times 3}$$

$$\tilde{\Sigma}_\epsilon = \tilde{\Theta}_{dis}'\Sigma_\varepsilon\tilde{\Theta}_{dis}, \quad (\Gamma_0 - \Gamma_f\tilde{\Theta}_1)^{-1} = \tilde{\Theta}_{dis} \quad (63)$$

In the empirical analysis we compare the QRE-NK-DSGE model with the standard NK counterpart model under RE, that is the model obtained with $\Phi_2 = 0_{3\times 3}$. The CER for the NK model under RE in eq.s (19) – (21).

5.1.1 Bayesian estimation

Table 4 summarizes the priors for the QR-NK-DSGE model. The priors for the structural parameters $\boldsymbol{\theta}$ are taken from Benati and Surico (2009) while the priors for additional parameters ϕ are centered on Rational Expectation solution. This means that for all of the ϕ_s parameters, the prior distribution specified is a $N(0, 0.25)$.[5]

Table 4. Priors

Parameter	Density	Mode	Standard Deviation
δ	*Inverse Gamma*	0.06	0.04
ϖ_f	*Beta*	0.25	0.20
α	*Beta*	0.75	0.20
ϱ	*Gamma*	0.05	0.01
λ_r	*Beta*	0.75	0.20
λ_π	*Gamma*	1.00	0.50
λ_y	*Gamma*	0.15	0.25
ρ_y	*Beta*	0.25	0.20
ρ_π	*Beta*	0.25	0.20
ρ_i	*Beta*	0.25	0.20
σ_y^2	*Inverse Gamma*	0.10	0.25
σ_π^2	*Inverse Gamma*	0.50	0.50
σ_i^2	*Inverse Gamma*	0.25	0.25
$\boldsymbol{\phi}_{i,j}$	*Normal*	0	0.25

Notes: the priors are taken by Benati and Surico (2009) except for the auxiliary parameters that are centered in the RE hypothesis with mean 0 and variance sufficient small to ensure the stability of the solution companion matrix.

The posteriors, reported in Table 5, are coputed using a standard Random-Walk Metropolis (RWM) proposed in Section 4.1. Analyzing the difference

[5] The difference in the prior distribution of δ with respect to Benati and Surico (2009) is only a matter of specification and the property that if $X \sim Gamma(k, \theta)$ then $\frac{1}{X} \sim InverseGamma(k, \theta^{-1})$

between the posteriors in the RE and in QRE we can notice that ϖ_f has a posterior mean equal to 0.453 with a 90% credible set of $[0.343, 0.602]$ in the RE model while it has a posterior mean equal to 0.843 with a credible set of $[0.721, 0.939]$ in QRE. This aspect indicates the higher weight attribuited to the first lag in the RE model in the intertemporal IS curve (eq. (7)). Indeed $(1 - \varpi_f)$ is equal to $1 - 0.453 = 0.547$ in the model based on RE and equal to $1 - 0.843 = 0.157$ in the model based on the QRE. Looking now at $\phi_{1,2}$, the coefficient related to the second lag in the intertemporal IS curve (eq. (7)), we can see that it has mean equal to 0.256 with a credible set $[0.127, 0.382]$ that reflects the weight attributited at the first lag in the RE model (0.547) is actually divided in 0.157 for the first lag and 0.256 for the second lag in QRE. Another important difference is related to the coefficient ρ_π whose mean is 0.119 with a 90% credible set equal to $[0.101, 0.153]$ under RE and 0.800 with a 90% credible set equal $[0.671, 0.908]$ under QRE. The higer value of the autoregressive persistence parameter in the QRE model is mitigated by the coefficient $\phi_{2,2} = -0.721$, the parameter related to the second lag in the New Keynesian Phillips curve (eq. (59))

Table 5. Posteriors

	Posterior RE	Posterior QRE
Parameters	$Mean[5\%, 95\%]$	$Mean[5\%, 95\%]$
δ	$0.185[0.158, 0.199]$	$0.186[0.159, 0.199]$
ϖ_f	$0.453[0.343, 0.602]$	$0.843[0.721, 0.939]$
α	$0.058[0.037, 0.091]$	$0.061[0.037, 0.094]$
ϱ	$0.039[0.028, 0.053]$	$0.054[0.038, 0.071]$
λ_r	$0.756[0.675, 0.827]$	$0.744[0.659, 0.815]$
λ_π	$1.907[1.906, 1.908]$	$1.839[1.661, 2.087]$
λ_y	$0.775[0.382, 1.277]$	$1.112[0.778, 1.445]$
ρ_y	$0.439[0.266, 0.605]$	$0.825[0.713, 0.903]$
ρ_π	$0.119[0.101, 0.153]$	$0.800[0.671, 0.908]$
ρ_i	$0.592[0.272, 0.815]$	$0.832[0.748, 0.902]$
σ_y^2	$0.073[0.047, 0.106]$	$0.135[0.103, 0.173]$
σ_π^2	$0.245[0.202, 0.297]$	$0.247[0.204, 0.297]$
σ_i^2	$0.117[0.096, 0.142]$	$0.121[0.099, 0.146]$
$\phi_{1,2}$	$-$	$0.256[0.127, 0.382]$
$\phi_{2,2}$	$-$	$-0.721[-0.957, -0.470]$
$\phi_{3,2}$	$-$	$0.174[0.012, 0.369]$

Notes: the posteriors are obtained using a Random-Walk Metropolis (RWM) algorithm. The posteriors satisfy the standard convergence criteria and the acceptance ratio is 21.47% for the RE and 27.04% for the QRE. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

In Table 6 we report the DIC (Deviance Information Criterion) and PMCC (Predictive Model Choice Criterion) criteria described in Section 4.1. These criteria stress the better time series performance of the QRE with respet to the RE. Indeed in both cases, the DIV and PMCC of the model under QRE are less (79.79 vs 113.52 for the Predictive Model Choice Criteria and 107.72 vs 212.97 for Deviance Information Criteria).

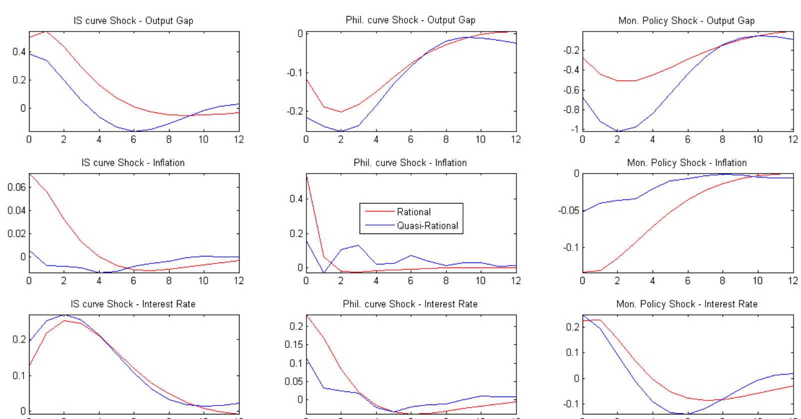
Table 6. Goodness of fit

	DIC	*PMCC*
RE	212.97	113.52
QRE	107.72	79.17

Notes: Deviance Information Criterion (DIC) and Predictive Model Chioce Criteria (PMCC) proposed in Section 4.1. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

Finally, Figure 2 plots the Impulse Response Functions (IRFs) in the two different formulations. The IRFs under RE have a relatively standard pattern, in the sense that they are similar to the IRFs documented in Benati and Surico (2009). The IRFs computed based on QR-NK model are similar, the main difference is the response of the inflation given an inflation shock, which reflect the already discussed role of the $\phi_{2,2}$ coefficient.

Figure 2. Impulse Response Functions

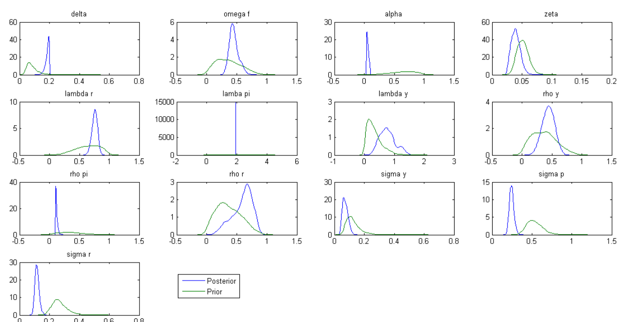


Notes: Impulse Response Functions are computed under RE and QRE expectations. In the first column the responses to an aggregate demad shock, in the second column the responses to inflation shock and in the last column the responses to a monetary policy shock. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

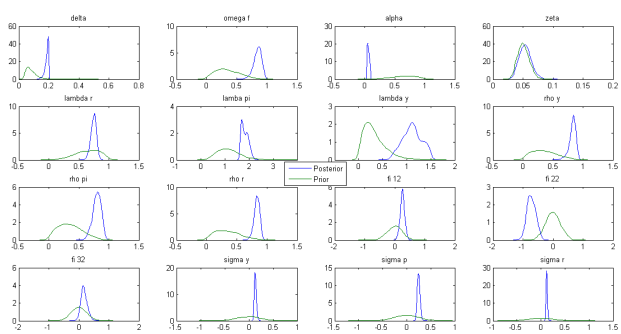
The last analysis is the comparison between prior and posterior distributions. Figure 3 reports the graphs.

Figure 3. Priors and Posteriors

Rational



Quasi-Rational



5.1.2 Frequentist estimation

In this section we estimate and evaluate the QR-NK model with a frequentist estimation using the algorithm proposed in Section 4.2. Table 7 reports the estimated parameters and Table 8 some goodness of fit criteria.

Table 7. Estimations

Parameters	$RE(std.err.)$	$QRE(std.err.)$
δ	0.200(0.182)	0.200(0.057)
ϖ_f	0.708(0.078)	0.848(0.384)
α	0.035(0.190)	0.035(0.012)
ϱ	0.025(0.012)	0.025(0.008)
λ_r	0.889(0.057)	0.961(0.286)
λ_π	1.650(1.157)	2.532(0.871)
λ_y	0.419(0.201)	1.021(0.518)
ρ_y	0.857(0.044)	0.969(0.076)
ρ_π	0.710(0.074)	0.839(0.302)
ρ_i	0.489(0.098)	0.611(0.374)
σ_y^2	0.152(0.032)	0.166(0.056)
σ_π^2	0.043(0.010)	0.060(0.058)
σ_i^2	0.010(0.002)	0.007(0.006)
$\phi_{1,2}$	$-$	0.086(0.031)
$\phi_{2,2}$	$-$	$-0.545(0.286)$
$\phi_{3,2}$	$-$	$-0.054(0.018)$

Notes: The log-likelihood is maximized by a Kalman-filtering approach and the simulated-annealing/genetic algorithm of Andreasen (2010), using the following bounds for the parameters: [0.010-0.200] for δ; [0.100-0.999] for ϖ_f ; [0.035-0.100] for α; [0.025-8] for ϱ; [0.001-0.999] for λ_r; [0.001-1.500] for λ_y; [1.650-5.500] for λ_π; [0.001-0.999] for ρ_y, ρ_π and ρ_i, and leaving all remaining parameters (including the auxiliary parameters $\phi_{i,j}, i = 1, 2, 3$ and $j = 2$ collected in the vector ϕ) free on condition that model's solution uniqueness and stability is satisfied. Different initial values have been used for $\theta^* = (\theta', \phi')'$ converging always to the same maximum. Standard errors in parentheses [have been calculated from the Hessian matrix using hessian function in matlab]. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

At first glance, we notice that the auxiliary parameters $\phi = \left(\phi_{1,2}, \phi_{2,2}, \phi_{3,2}\right)' = dg\left(\Phi_2\right)$ estimated under QRE, reported in the second coloum of Table 7, are all significant. The main differences characterize the coefficients of the Taylor rule. In practical the coefficient λ_π indicating the long-run response of the Central Bank to inflation, is equal to 1.650 in the model under RE and to 2.532 in the model estimated under QRE. Analyzing λ_y, the long-run response of the Central Bank to output gap, we can observe that this coefficient is equal to 0.419 under RE and to 1.021 under QRE. The higher value of this two parameters indicate an higher reaction of the monetary policy to the inflation and to the output gap.

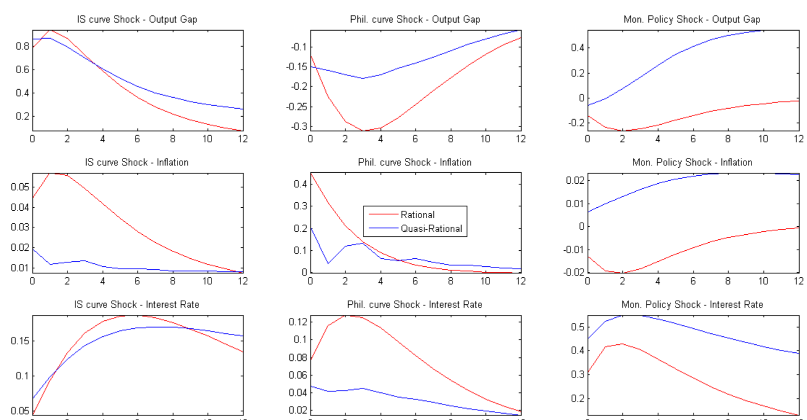
Table 8. Goodness of fit

	RE	QRE
Likelihood	30.88	59.11
Akaike	-35.77	-86.22^*
Hannan-Quinn	-22.18	-69.63^*
Schwart	-2.16	-45.19^*
LR(RE vs QRE)=40.46, p-value<0.0001		
LR(CER under RE)=32.70, p-value<0.0001		
LR(CER under QRE)=55.42, p-value<0.0001		

Notes: the LR(RE vs QRE) is computed comparing the log-likelihood obtained with RE and the log-likelihood with QRE. The LR(CER under RE) is computed comparing the log-likelihood with RE and the log-likelihood with unrestricted VAR(2). The LR(CER under QRE) is computed comparing the log-likelihoods with QRE and the log-likelihood with unrestricted VAR(3). Akaike, Hannan-Quinn and Schwarz criteria are described in Section 4.1. Asterisks denote the optimal lag selection according to the information criterion. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

The likelihood ratio test and the goodness of fit criteria in Table 8 show the better time series performance of the model under QRE. All chosen criteria select QRE as the optimum model, -35.77 vs -86.22 for Akaike criterion, -22.18 vs -69.63 for Hannan and Quinn criterion and -2.16 vs -45.19 for Schwarz criterion. Also the LR test reject the null model (RE) with a $p-value < 0.0001$ (in this case the LR(RE vs QRE) $\sim \chi_3$). In Figure 4 the IRFs are plotted for the two cases. The main differences are shown by the responses of the ountput gap and of the inflation to a monetary policy shock. In this two cases QRE and RE report an opposite response.

Figure 4. Impulse Response Functions

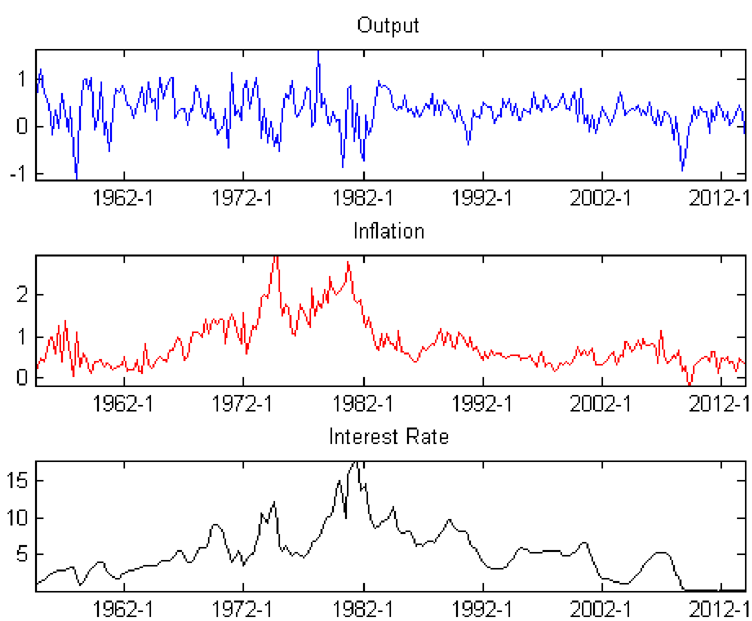


Notes: Impulse Response Functions are computed under RE and QRE expectations. In the first column the responses to an aggregate demad shock, in the second column the responses to inflation shock and in the last column the responses to a monetary policy shock. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (58)-(61).

5.2 Unobservable case

In this section we estimate the QR-NK-DSGE model for the case where the variable $\tilde{y}_t$ is assumed to be not directly unobservable. The reference model is proposed in Section 1.2, case B. Differently from the previous case now the vector of observed variables is $Y_t = [\Delta y_t, \pi_t, i_t]'$ and the measurement system is in eq.s $(13) - (15)$.

Figure 5. Time Series



As in the previous application, the discount factor parameter β is fixed to 0.99 (corresponding to an annual discount rate of approximately 4%). Again, we drop out from the analysis the period before the Volcker stabilization (1979) and after the financial crisis (2008:Q4) and we focus on the period 1984:Q2-2008:Q3 ($T = 98$). As described in Section 3 the representation of a DSGE in this case is a state space model as in eq.s $(32) - (33)$. In this case, the 'best fitting' model for Y_t is based on $k = 4$ lags

$$s.e. \quad : \begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-k+1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ I_3 & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_3 & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & I_3 & 0_{3\times3} \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-k} \end{bmatrix} + \begin{bmatrix} A_{dis} \\ 0_{n\times n} \\ 0_{n\times n} \\ 0_{n\times n} \end{bmatrix} \epsilon_t \tag{64}$$

$$m.e. \quad : \begin{bmatrix} \Delta y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-k+1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \epsilon_{y,t} \tag{65}$$

Where $A_i, i = 1, ..., 4, dis$ are 3×3 matrices whose only restriction is to fulfil the minimality (controllability and observability) and local identification conditions discussed in Komunjer and Ng (2011). Table 9 reports the information criteria for lag the length selection. The LR test selects $k = 4$ lags while the information criteria choose different dynamics: Akaike selects

$k = 5$ lags, Hannan and Quinn $k = 2$ lags Schwarz $k = 2$ lags

Table 9. Lag lenght selection

lag	Likelihood	LR	p-value	Akaike	Hannan-Quinn	Schwarz
2	151.42	74.71	< 0.0001	-240.84	-208.57^{*}	-161.03^{*}
3	161.08	55.40	0.001	-242.15	-200.69	-139.58
4	176.99	23.57	0.167^{*}	-255.98	-205.42	-130.84
5	186.44	4.68	0.861	-256.87^{*}	-197.29	-109.36
6	188.78	$-$	$-$	-243.55	-175.04	-73.87

Notes: the log-likelihood is maximized by a Kalman-filtering approach and the simulated-annealing/genetic algorithm of Andreasen's (2010). The LR tests are computed by comparing the log-likelihoods obtained with $k = 2, ..., 6 = k$ lags and the log-likelihood obtained with $k = 6$ lags. Asterisks denote the optimal lag selected by the test/criterion. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (64)-(65).

Following the indication of the LR test we choose $k = 4$ lags. Given $k = 4$, from Definition 2 the QR-NK-DSGE pseudo structural model is

$$\begin{aligned}
\Gamma_0 Z_t &= \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \Phi_2 Z_{t-2} + \Phi_3 Z_{t-3} + \upsilon_t \\
\upsilon_t &= R\upsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN\left(0, \Sigma_\varepsilon\right)
\end{aligned}$$

where the matrices Γ_0, Γ_f and Γ_b are described in eq.s $(11) - (12)$ and

$$\Phi_2 = \begin{bmatrix} \phi_{1,2} & 0 & 0 \\ 0 & \phi_{2,2} & 0 \\ 0 & 0 & \phi_{3,2} \end{bmatrix}, \Phi_2 = \begin{bmatrix} \phi_{1,3} & 0 & 0 \\ 0 & \phi_{2,3} & 0 \\ 0 & 0 & \phi_{3,3} \end{bmatrix}, R = \begin{bmatrix} \rho_y & 0 & 0 \\ 0 & \rho_\pi & 0 \\ 0 & 0 & \rho_i \end{bmatrix}.$$

In detail, the pseudo-structural model is given by

$$\begin{aligned}
\tilde{y}_t &= \varpi_f E_t \tilde{y}_{t+1} + (1-\varpi_f)\tilde{y}_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \phi_{1,2}\tilde{y}_{t-2} + \phi_{1,3}\tilde{y}_{t-3} + \upsilon_{y,t} \qquad (66)\\
\pi_t &= \frac{\beta}{1+\alpha\beta} E_t \pi_{t+1} + \frac{\alpha}{1+\alpha\beta}\pi_{t-1} + \varrho \tilde{y}_t + \phi_{2,2}\pi_{t-2} + \phi_{2,3}\pi_{t-3} + \upsilon_{\pi,t}\\
i_t &= \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_t + \lambda_y \tilde{y}_t) + \phi_{3,2} i_{t-2} + \phi_{3,3} i_{t-3} + \upsilon_{i,t}\\
\upsilon_t^x &= R\upsilon_{t-1}^x + \varepsilon_{x,t}, \quad \varepsilon_{x,t} \sim WN\left(0, \sigma_x^2\right), \quad x = \tilde{y}, \pi, i \qquad (67)
\end{aligned}$$

and the measurement equation in eq.s (13) – (15) takes the form

$$\begin{bmatrix} \Delta y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ \pi_t \\ i_t \\ \tilde{y}_{t-1} \\ \pi_{t-1} \\ i_{t-1} \\ \tilde{y}_{t-2} \\ \pi_{t-2} \\ i_{t-2} \\ \tilde{y}_{t-2} \\ \pi_{t-2} \\ i_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tau_{y,t}. \qquad (68)$$

The Step 2 of the estimation procedure summarized in Section 4 requires estimating $\theta^{*}=(\theta',\phi')'$ from the model in eq.s (64) – (65) by imposing the CER $A_j=\tilde{\Theta}_j$, $j=1,2,3,4$ and $\Sigma_\epsilon=\tilde{\Sigma}_\epsilon$, where

$$\left(\Gamma_0+R\Gamma_f\right)\tilde{\Theta}_1-\Gamma_f\left(\tilde{\Theta}_1^2+\tilde{\Theta}_2\right)-\left(\Gamma_b+R\Gamma_0\right)=0_{3\times3} \tag{69}$$

$$\left(\Gamma_0+R\Gamma_f\right)\tilde{\Theta}_2-\Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_2+\tilde{\Theta}_3\right)-\left(\Phi_2-R\Gamma_b\right)=0_{3\times3}$$

$$\left(\Gamma_0+R\Gamma_f\right)\tilde{\Theta}_3-\Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_3+\tilde{\Theta}_4\right)-\left(\Phi_3-R\Phi_2\right)=0_{3\times3}$$

$$\left(\Gamma_0+R\Gamma_f\right)\tilde{\Theta}_4-\Gamma_f\tilde{\Theta}_1\tilde{\Theta}_4+R\Phi_3=0_{3\times3}$$

$$\tilde{\Sigma}_\epsilon=\tilde{\Theta}_{dis}'\Sigma_\varepsilon\tilde{\Theta}_{dis},\quad \left(\Gamma_0-\Gamma_f\tilde{\Theta}_1\right)^{-1}=\tilde{\Theta}_{dis} \tag{70}$$

The counterpart NK model under RE is given by

$$\begin{aligned}\Gamma_0 Z_t &= \Gamma_f E_t Z_{t+1}+\Gamma_b Z_{t-1}+\upsilon_t\\ \upsilon_t &= R\upsilon_{t-1}+\varepsilon_t, \varepsilon_t\sim WN\left(0,\Sigma_\varepsilon\right)\end{aligned}$$

where

$$R=\begin{bmatrix}\rho_y & 0 & 0\\ 0 & \rho_\pi & 0\\ 0 & 0 & \rho_i\end{bmatrix}.$$

with the follow measurement equation

$$\begin{bmatrix} \Delta y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ \pi_t \\ i_t \\ \tilde{y}_{t-1} \\ \pi_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tau_{y,t}. \tag{71}$$

The CER for the standard NK model are summarized in eq.s $(34)-(35)$ (the CER are the same as in the system $(69)-(70)$ with $\Phi_2 = \Phi_3 = 0_{3\times 3}$).

5.2.1 Bayesian estimation

Table 10 summerizes the priors used for the QR-NK-DSGE model. As for the observable case, the priors for the truly structural parameters $\boldsymbol{\theta}$ are taken from Benati and Surico (2009), while the priors for the additional parameters $\boldsymbol{\phi}$ are centered in Rational Expectation solution. Hence for all $\phi_{i,j}$, it is used a $N(0, 0.25)$ distribution. For the variance of the measurement error, σ_τ^2, the prior is the same as for σ_y^2.

Table 10. Priors

Parameter	Density	Mode	Standard Deviation
δ	*Inverse Gamma*	0.06	0.04
ϖ_f	*Beta*	0.25	0.20
α	*Beta*	0.75	0.20
ϱ	*Gamma*	0.05	0.01
λ_r	*Beta*	0.75	0.20
λ_π	*Gamma*	1.00	0.50
λ_y	*Gamma*	0.15	0.25
ρ_y	*Beta*	0.25	0.20
ρ_π	*Beta*	0.25	0.20
ρ_i	*Beta*	0.25	0.20
σ_y^2	*Inverse Gamma*	0.25	0.25
σ_π^2	*Inverse Gamma*	0.50	0.50
σ_i^2	*Inverse Gamma*	0.25	0.25
σ_τ^2	*Inverse Gamma*	0.25	0.25
$\boldsymbol{\phi}_{i,j}$	*Normal*	0	0.25

Notes: the priors are taken by Benati and Surico (2009) except for the auxiliary parameters that are centered in the RE hypothesis with mean 0 and variance sufficient small to ensure the stability of the solution companion matrix.

The posteriors, reported in Table 11, are computed using a standard Random-Walk Metropolis (RWM) discussed in Section 4.1. Analyzing the posterios it's possible to observe that the main differences are in the coeffients of the monetary policy rule λ_π and λ_y. In particoular λ_π, the long-run response of the Central Bank to inflation, switches from 2.107 with 90% credible set of $[1.682, 3.043]$ under RE to 1.801 with a credible set $[1.658, 2.120]$ under QRE. The parameter λ_y, the long-run response of the Central Bank to the output gap, passes from 0.449 with a credible set $[0.116, 0.805]$ under RE to 1.054 with a credible set $[0.167, 1.478]$ under QRE. The difference in this two parameters indicates that the additional lags in the structural model shrink the gap between the response of Central Bank to inflation and the response of Central Bank to the output gap. Another difference is given by the coefficient $\rho_\pi = 0.484$ with a credible set $[0.175, 0.771]$ under RE and 0.783 with a credible set $[0.637, 0.903]$ under QRE. This difference is explained by the coefficient $\phi_{2,2} = -0.638$ with a credible set $[-0.913, -0.370]$ that mitigates the higher value of ρ_π in QRE with respect to RE.

Table 11. Posteriors

	Posterior RE	Posterior QRE
Parameters	$Mean[5\%, 95\%]$	$Mean[5\%, 95\%]$
δ	$0.183[0.156, 0.199]$	$0.185[0.158, 0.199]$
ϖ_f	$0.136[0.102, 0.200]$	$0.829[0.649, 0.951]$
α	$0.056[0.036, 0.088]$	$0.062[0.037, 0.093]$
ϱ	$0.053[0.038, 0.072]$	$0.053[0.038.0.071]$
λ_r	$0.783[0.682, 0.870]$	$0.733[0.584, 0.902]$
λ_π	$2.107[1.682, 3.043]$	$1.801[1.658, 2.120]$
λ_y	$0.449[0.116, 0.805]$	$1.054[0.167, 1.478]$
ρ_y	$0.529[0.341, 0.714]$	$0.845[0.681, 0.952]$
ρ_π	$0.484[0.175, 0.771]$	$0.783[0.637, 0.903]$
ρ_i	$0.470[0.204, 0.717]$	$0.641[0.233, 0.916]$
σ_y^2	$0.041[0.030, 0.054]$	$0.047[0.034, 0.063]$
σ_π^2	$0.232[0.193, 0.278]$	$0.249[0.204, 0.301]$
σ_i^2	$0.112[0.092, 0.135]$	$0.115[0.094, 0.139]$
σ_τ^2	$0.048[0.035, 0.063]$	$0.054[0.039, 0.073]$
$\phi_{1,2}$	$-$	$-0.082[-0.364, 0.209]$
$\phi_{2,2}$	$-$	$-0.638[-0.913, -0.370]$
$\phi_{3,2}$	$-$	$0.125[-0.176, 0.439]$
$\phi_{1,3}$	$-$	$0.011[-0.249, 0.263]$
$\phi_{2,3}$	$-$	$-0.247[-0.538, 0.055]$
$\phi_{3,3}$	$-$	$-0.053[-0.307, 0.213]$

Notes: the posteriors are obtained using a Random-Walk Metropolis (RWM) algorithm. The posteriors satisfy the standard convergence criteria and the acceptance ratio is 22.94% for the RE and 36.79% for the QRE. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

The DIC criterion in Table 12 shows the better time series performance of the QR-NK-DSGE model compared to RE model.

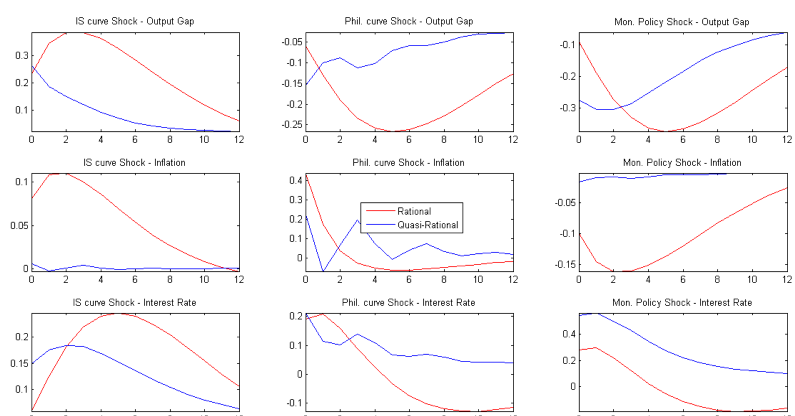
Table 12. Goodness of Fit

	DIC
RE	42.41
QRE	3.26

Notes: Deviance Information Criterion (DIC) proposed in Section 4.1. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

In Figure 6 we plot the IRFs for RE and QRE. The IRFs under RE and QRE are very similar.

Figure 6. Impulse Response Functions

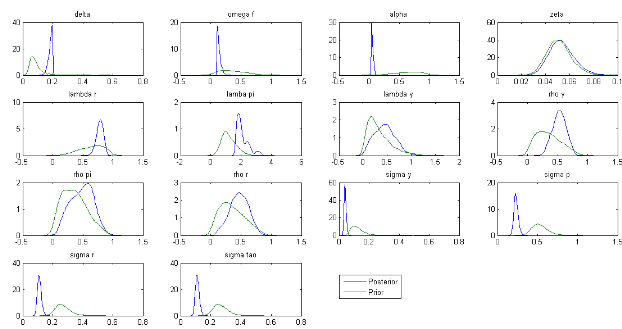


Notes: Impulse Response Functions are computed under RE and QRE expectations. In the first column the responses to an aggregate demad shock, in the second column the responses to inflation shock and in the last column the responses to a monetary policy shock. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

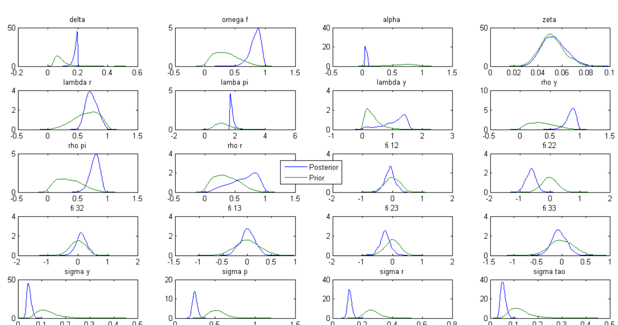
Figure 7 reports the comparison between prior and posterior distributions

Figure 7. Priors and Posteriors

Rational



Quasi-Rational



5.2.2 Frequentist estimation

The estimations in Table 13 are obtained using the algorithm in Section 4.2.

Table 13. Estimations

Parameters	$RE(std.err.)$	$QRE(std.err.)$
δ	0.010(0.057)	0.079(0.055)
ϖ_f	0.572(0.062)	0.269(0.207)
α	0.035(0.230)	0.035(0.039)
ϱ	0.041(0.121)	0.027(0.043)
λ_r	0.908(0.054)	0.889(0.034)
λ_π	1.650(0.974)	1.650(0.803)
λ_y	0.336(0.963)	1.500(0.248)
ρ_y	0.908(0.034)	0.801(0.190)
ρ_π	0.100(0.342)	0.775(0.082)
ρ_i	0.539(0.080)	0.192(0.157)
σ_y^2	0.001(0.001)	0.006(0.002)
σ_π^2	0.025(0.003)	0.053(0.010)
σ_i^2	0.011(0.002)	0.006(0.001)
σ_τ^2	0.045(0.009)	0.031(0.006)
$\phi_{1,2}$	$-$	$-0.061(0.191)$
$\phi_{2,2}$	$-$	$-0.444(0.176)$
$\phi_{3,2}$	$-$	0.057(0.061)
$\phi_{1,3}$	$-$	0.047(0.016)
$\phi_{2,3}$	$-$	0.065(0.131)
$\phi_{3,3}$	$-$	$-0.192(0.058)$

Notes: The log-likelihood is maximized by a Kalman-filtering approach and the simulated-annealing/genetic algorithm of Andreasen (2010), using the following bounds for the parameters: [0.010-0.200] for δ; [0.100-0.999] for ϖ_f ; [0.035-0.100] for α; [0.025-8] for ϱ; [0.001-0.999] for λ_r; [0.001-1.500] for λ_y; [1.650-5.500] for λ_π; [0.001-0.999] for ρ_y, ρ_π and ρ_i, and leaving all remaining parameters (including the auxiliary parameters $\phi_{i,j}, i = 1,2,3$ and $j = 2$ collected in the vector ϕ) free on condition that model's solution uniqueness and stability is satisfied. Different initial values have been used for $\theta^* = (\theta', \phi')'$ converging always to the same maximum. Standard errors in parentheses [have been calculated from the Hessian matrix using hessian function in matlab].The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

We notice at first that the auxiliary parameters $\phi_{2,2}, \phi_{1,3}$ and $\phi_{3,3}$ estimated under QRE in the second coloum of Table 13 are statistically significant. This is a first evidence in support of QRE. Secondly, the main differences in the estimations regard the parameters of the New-Keynesian Phillips Curve (henceforth NK-PC, eq. (8)). The parameter ϱ, the slope of the NK-PC, is quite larger under RE relative QRE (0.041 to 0.027) indicating a more flat-sloped NK-PC under QRE. Another diffence is in the autoregressive persistence parameter, ρ_π, that is 0.100 under RE and 0.775 under QRE. This sharp difference can be explained by the additional lag $\phi_{2,2} = -0.444(0.176)$ of the NK-PC. The last difference we want to underline is the parameter λ_y, the long-run coefficient on the output gap in the monetary policy rule, that goes from 0.336(0.963) in the RE model to 1.500(0.248) in the QRE model indicating an higher reaction of the policy rate to movements in the output gap.

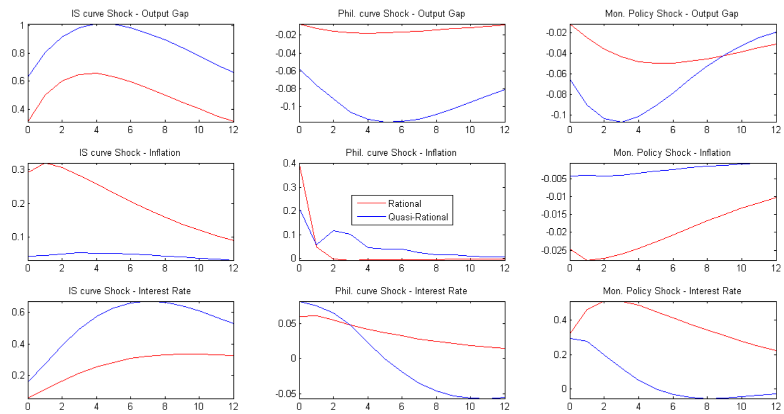
Table 14. Goodness of fit

	RE	**QRE**
Likelihood	115.09	129.78
Akaike	-202.18	-219.58^*
Hannan-Quinn	-187.54	-198.66^*
Schwarz	-165.99	-167.87^*
LR(RE vs QRE)=29.40, p-value<0.0001		
LR(CER under RE)=72.66, p-value<0.0001		
LR(CER under QRE)=94.40, p-value<0.0001		

Notes: the LR(RE vs QRE) is computed comparing the log-likelihood obtained with RE and the log-likelihood with QRE. The LR(CER under RE) is computed comparing the log-likelihood with RE and the log-likelihood with unrestricted state-space with 2 lags. The LR(CER under QRE) is computed comparing the log-likelihoods with QRE and the log-likelihood with unrestricted state-space with 4 lags. Akaike, Hannan-Quinn and Schwarz criteria are described in Section 4.1. Asterisks denote the optimal lag selection according to the information criterion. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

From Table 14 we can notice that according to each criterion the selected model is the QRE, -219.58 vs -202.18 for the Akaike criterion, -2198.66 vs -187.54 for the Hannan-Quinn criterion and -167.88 vs -165.99 for the Schwarz criterion. Also the likelihood ratio test strongly select the QRE. Figure 8 shows the Impulse Response Functions. The responses under RE hypothesis are quite standard and in line with the ones porposed in Benati and Surico (2009). On the other hand the IRFs in the QRE model show more persistent but there are no big differences between the shape of the IRFs in the two different models.

Figure 8. Impulse Response Functions



Notes: Impulse Response Functions are computed under RE and QRE expectations. In the first column the responses to an aggregate demad shock, in the second column the responses to inflation shock and in the last column the responses to a monetary policy shock. The sample is 1984:Q2-2008:Q3 and the model is described in eq.s (66)-(67) with measurement equation in eq. (68).

6 Conclusion

The poor time-series performance of this class of models, which dominate the New Keynesian macroeconomic tradition, is generally ascribed to the tight nature of the restrictions these models impose on state-space/VAR representation for the data. The restrictions NK-DSGE place on state-space/VAR models can be classified into two categories: (i) highly nonlinear cross-equation restrictions (CER) which the system places on its unique stable reduced form solution and which can be potentially used to recover estimates of the structural parameters; (ii) constraints on the lag order of the reduced form solution. This work find a possible solutions to (ii) type of misspecification.

To try to solve this problem, a growing literature attempts to 'take DSGE models to the data'. In this work we focus on small-scale NK-DSGE models used in monetary policy and business cycle analysis on quarterly data and propose a solution that allows one to reconcile the gap between the predictions of theory and the autocorrelation structure of the data.

Under RE, the reduced form solution of NK-DSGE models gives also rise to (implicit) zero restrictions that affect the actual autocorrelation structure of the data. To avoid the misspecification that typically affects these models, the probabilistic structure of the data has been usually completed manipulating arbitrarily the shock structure of the model or using prior distributions for the parameters, with the possibility of relaxing the CER. This idea of selecting rich processes for the error terms is very common in literature but it is completely aribtrary and it is in contrast with models (DSGE) based on strong economic principles. With this work we define a data-driven process able to define the 'true' dynamic of the data and use that information to define a pseudo-structural model whose reduced form solution capture the whole dynamics in the data. We adapt the dynamic specification of the NK-DSGE model such that the CER does not include zero restrictions. A central role in this work is taken by the 'best fitting' unrestricted model for the data as driver for dynamic specification. We denote our approach Quasi-Rational Expectations because while not renouncing to the concept of model-consistent expectations, the starting point of our analysis is a model for the data, not the structural form: the estimable structural form is obtained in a second step such that there does not exist any mismatch between the autocorrelation structure featured by the two models. In this way we can circumvent the omitted dynamics issue embodied by NK-DSGE models. Quasi Rational DSGE is defined as a linear rational expectations model derived from the baseline structural specification, such that its stable reduced form solution has the same lag structure as the state space (VAR) which fits the data optimally. The advantage in this framework is that one does not need to dismiss a-priori the theoretical foundations upon which DSGE models are built on but at the same time it is possible to model the business cycle facts by relaxing the tight assumption that apart from structural parameters the agents (as well as the econometrician) know the underlying data generating process.

In this research we analyze the observable and unobservable case and for each application we propose both bayesian and frequentist estimation. In this way we cover a very wide range of possibilities under which DSGE models are usually analyzed. In literature the debate on using proxies for unobservable components or computing these variables from the system is still open hence in this work we define the pseudo-structural QR-NK-DSGE model for the two different cases. Also the debate between Bayesian and frequentis estimation is still open even if in the recent year the bayesian applications in DSGE framework occupy most of the scientific papers. However, the last case analyzed in the thesis (presence of unobservable components and frequentist estimation) is the most challenging. Indeed in this case identification and optimization problems need be solved. Identification issue is a very challenging aspect in DSGE evaluation because e proper investigation of this problem requires a clarification on the relationship existing between DSGE models and state-space/VAR representations for the observed variables.

Anyway, applications to the US economy provide reliable estimates of the structural parameters and show a better time series performance of the QR-NK-DSGE with respect to the standard NK-DSGE model in all case considered.

7 References

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

Andreasen, M. M. (2010), How to maximize the likelihood function for a DSGE model.Computational Economics 35.2 (2010): 127-154.

Bardsen, G., Fanelli, L. (2014), Frequentist evaluation of small DSGE models, Journal of Business Economic and Statistics. Forthcoming.

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

Benati, L., Surico, P. (2009). VAR Analysis and the Great Moderation. American Economic Review, 99(4), 1636-52.

Binder, M., Hashem M. P. (1995), Multivariate rational expectations models and macroeconomic modelling: a review and some new results. No. 9415. Faculty of Economics, University of Cambridge.

Bovin, 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.

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

Canova, F., Ferroni, F. (2012), The dynamics of US inflation: Can monetary policy explain the changes? Journal of Econometrics 167, 47-60.

Castelnuovo, E. and Fanelli, L., (2014), Monetary policy indeterminacy and identification failures in the U.S.: results from a robust test, Journal of Applied Econometrics, forthcoming.

Cavaliere, G., Rahbek, A., Taylor, M.R. (2012), Bootstrap determination of the cointegration rank in vector autoregressive models, Econometrica 80(4), 1721-1740

Chevillon, G., Massmann, M., Mavroeidis, S. (2010), Inference in models with adaptive learning, Journal of Monetary Economics 57, 341-351.

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, Journal of Economic Dynamics and Control 25, 1485-1525.

Cole, S. J., Milani, F. (2014), The Misspecification of Expectations in New Keynesian Models: A DSGE-VAR Approach. Working paper.

Consolo, A., Favero, C.A., Paccagnini, A. (2009), On the statistical identification of DSGE models, Journal of Econometrics 150, 99-115.

Cúrdia,V., Reis, R. (2010), Correlated disturbances and U.S. business cycle, Federal Reserve Bank of New York, Staff Report No. 434.

Dees, S., Pesaran, H.M., Smith, V., Smith, R.P. (2008), Identification of New Keynesian Phillips curves from a global perspective, ECB Working Paper No. 892.

DeJong, D.N., Dave, C. (2007), Structural Macroeconometrics, Princeton University Press, Princeton.

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 Analysis with Potentially Misspecified Models, American Economic Review 99(4), 1415-1450.

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.

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

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. (2009), Estimation of Quasi-Rational DSGE Monetary Models, Department of Statistical Sciences, Working Paper.

Fanelli, L., Palomba, G. (2011), Simulation-based tests of forward-looking models under VAR learning dynamics, Journal of Applied Econometrics 26, 762-782.

Fernandez-Villaverde, J., J. Rubio-Ramirez, T. J. Sargent and M. W. Watson (2007). ABCs (and Ds) of understanding VARs. American Economic Review, 97, 1021-1026.

Franchi, M., Paruolo, P. (2014). Minimality of State Space Solutions of DSGE Models and Existence Conditions for Their VAR Representation. Computational Economics, forthcoming.

Franchi, M., Vidotto, A. (2013), A check for finite order VAR representations of DSGE models, Economics Letters 120, 100-103.

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

Fukač, M., Pagan, A. (2009), Limited information estimation and evaluation of DSGE models, Journal of Applied Econometrics, 25, 55-70.

Gelfand, A. E., Ghosh, S. K. (1998), Model choice: a minimum posterior predictive loss approach. Biometrika 85.1 (1998): 1-11.

Gorodnichenko, Y., Ng, S. (2010), Estimation of DSGE models when the data are persistent, NBER Working Paper, No. 15187.

Guerron-Quintana, P., A. Inoue, Kilian, L. (2013), Frequentist inference in weakly identified dynamic stochastic general equilibrium models: Acronyms must be spelled out in titles for indexing purposes. Quantitative Economics 4.2 (2013): 197-229.

Hansen, N. (2005), The CMA evolution strategy: A tutorial. Vu le 29 (2005).

Holt, M.T., Mckenzie, A. M. (2003), Quasi-rational and ex ante price expectations in commodity supply models: an empirical analysis of the broiler market, Journal of Applied Econometrics 18, 407-426.

Iskrev, N. (2008), Evaluating the information matrix in linearized DSGE models, Economic Letters 99, 607-610.

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.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems. Journal of Fluids Engineering 82.1 (1960): 35-45.

Kapetanios, G., Pagan, A., Scott, A. (2007), Making a match: Combining theory and evidence in policy-oriented macroeconomic modelling, Journal of Econometrics 136, 565-594.

Klein, P. (2000), 'Using the generalized schur form to solve a multivariate linear rational expectations model', Journal of Economic Dynamics and Control 24.

Komunjer, I., Ng, S. (2011), Dynamic identification of dynamic stochastic general equilibrium models, Econometrica 79, 1995-2032.

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.

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

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

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

McConnell, M., G. Perez-Quiros (2000), Output Fluctuations in the United States: What Has Changed Since the Early 1980s?, American Economic Review, 90, 1464-1476.

Muth, J. F. (1961), Rational Expectations and 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. (1998), 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.

Pesaran H. M., Smith, R. P. (2011), Beyond the DSGE straitjacket, The Manchester School 79, 5-16.

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

Ravenna, F. (2007), Vector autoregressions and reduced form representations of DSGE models, Journal of Monetary Economics 54, 2048-2064.

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

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

Ruge-Murcia, F.J. (2007), Methods to estimate dynamic stochastic general equilibrium models, Journal of Economic Dynamics and Control 31, 2599-2636.

Sims, C. A., (2002), Solving Linear Rational Expectations Models, Computationl Economics, 20(1-2), 1-2.

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.

8 Appendix

8.1 Appendix A

Derivation of the Measurement Equation. Given $\tilde{y}_t = y_t - y_t^p$ and the assumption that y_t^p follows a Random Walk process

$$y_t^p = y_{t-1}^p + \tau_{y,t}, \quad \tau_{y,t} \sim WN\left(0, \sigma_\tau^2\right) \tag{72}$$

then

$$\begin{aligned} \tilde{y}_t &= y_t - y_t^p \\ \tilde{y}_t &= y_t - y_t^p + \tilde{y}_{t-1} - \tilde{y}_{t-1} \\ \tilde{y}_t - \tilde{y}_{t-1} &= y_t - y_t^p - \left(y_{t-1} - y_{t-1}^p\right) \\ \Delta y_t &= \Delta\tilde{y}_t + y_t^p - y_{t-1}^p \end{aligned}$$

now under the assumption in eq. (72) we obtain

$$\Delta y_t = \Delta\tilde{y}_t + \tau_{y,t},$$

■

8.2 Appendix B

CER under Rational Expectations hypothesis. Given the structural model in eq.s (16) – (17)

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \upsilon_t \quad (73)$$

$$\upsilon_t = R\upsilon_{t-1} + \varepsilon_t \quad (74)$$

Substituting eq. (74) in qe. (73) the model become

$$\begin{aligned}
\Gamma_0 Z_t &= \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + R\upsilon_{t-1} + \varepsilon_t \\
\Gamma_0 Z_t &= \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + R\left(\Gamma_0 Z_{t-1} - \Gamma_f\left(Z_t - \eta_t\right) - \Gamma_b Z_{t-2}\right) + \varepsilon_t \\
\left(\Gamma_0 + R\Gamma_f\right) Z_t &= \Gamma_f E_t Z_{t+1} + \left(\Gamma_b + R\Gamma_0\right) Z_{t-1} - R\Gamma_b Z_{t-2} + R\Gamma_f \eta_t + \varepsilon_t \\
\breve{\Gamma}_0 \dot{Z}_t &= \breve{\Gamma}_f E_t \dot{Z}_{t+1} + \breve{\Gamma}_b \dot{Z}_{t-1} + \breve{\varepsilon}_t \\
\breve{\Gamma}_0 &= \begin{bmatrix} \Gamma_0 + R\Gamma_f & 0_{n\times n} \\ 0_{n\times n} & I_n \end{bmatrix}, \breve{\Gamma}_f \begin{bmatrix} \Gamma_f & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} \\
\breve{\Gamma}_b &= \begin{bmatrix} \Gamma_b + R\Gamma_0 & -R\Gamma_b \\ I_n & 0_{n\times n} \end{bmatrix}, \breve{\varepsilon}_t = \begin{bmatrix} \varepsilon_t + R\Gamma_f \eta_t \\ 0_{n\times n} \end{bmatrix}.
\end{aligned}$$

Now the CER are given by (see Castelnuovo and Fanelli, 2014)

$$\breve{\Gamma}_f \breve{A}^2 - \breve{\Gamma}_0 \breve{A} + \breve{\Gamma}_b = 0_{kn\times kn}$$

where $\breve{A}$ is the restricted companion matrix of the reduced solution

$$\breve{A} = \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix}$$

hence the CER become

$$\begin{bmatrix} \Gamma_f & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} \left(\begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix} \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix} \right)$$

$$- \begin{bmatrix} \Gamma_0 + R\Gamma_f & 0_{n\times n} \\ 0_{n\times n} & I_n \end{bmatrix} \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 \\ I_n & 0_{n\times n} \end{bmatrix} + \begin{bmatrix} \Gamma_b + R\Gamma_0 & -R\Gamma_b \\ I_n & 0_{n\times n} \end{bmatrix} = \begin{bmatrix} 0_{n\times n} & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix}$$

$$\begin{cases} (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_1 - \Gamma_f \left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2 \right) - (\Gamma_b + R\Gamma_0) = 0_{n\times n} \\ (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_2 - \Gamma_f \tilde{\Theta}_1 \tilde{\Theta}_2 + R\Gamma_b = 0_{n\times n} \end{cases}$$

■

8.3 Appendix C

CER under Quasi-Rational Expectations paradigm. Given the structural model in eq.s (28) – (29)

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \left(\sum_{j=2}^{k-1} \Phi_j Z_{t-j} \right) \mathbb{I}_{\{k\geq 3\}} + \upsilon_t^* \qquad (75)$$

$$\upsilon_t^* = R\upsilon_{t-1}^* + \varepsilon_t \qquad (76)$$

Substituting eq. (76) in qe. (75) and using some algebra the model become

$$\breve{\Gamma}_0 \dot{Z}_t = \breve{\Gamma}_f E_t \dot{Z}_{t+1} + \breve{\Gamma}_b \dot{Z}_{t-1} + \breve{\varepsilon}_t$$

$$\breve{\Gamma}_0 = \begin{bmatrix} \Gamma_0 + R\Gamma_f & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ 0_{n\times n} & I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} & I_n & \ddots & 0_{n\times n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0_{n\times n} & 0_{n\times n} & \cdots & \cdots & I_n \end{bmatrix}, \breve{\Gamma}_f = \begin{bmatrix} \Gamma_f & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \ddots & 0_{n\times n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \end{bmatrix}$$

$$\breve{\Gamma}_b = \begin{bmatrix} \Gamma_b + R\Gamma_0 & \Phi_2 - R\Gamma_b & \Phi_3 - R\Phi_2 & \cdots & -R\Phi_{k-1} \\ I_n & 0_{n\times n} & 0_{n\times n} & \cdots & 0_{n\times n} \\ 0_{n\times n} & I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{n\times n} & 0_{n\times n} & 0_{n\times n} & I_n & 0_{n\times n} \end{bmatrix}, \breve{\varepsilon}_t = \begin{bmatrix} \varepsilon_t + R\Gamma_f \eta_t \\ 0_{n\times n} \\ 0_{n\times n} \\ \vdots \\ 0_{n\times n} \end{bmatrix}.$$

Now the CER are given by (see Castelnuovo and Fanelli, 2014)

$$\breve{\Gamma}_f \breve{A}^2 - \breve{\Gamma}_0 \breve{A} + \breve{\Gamma}_b = 0_{kn\times kn}$$

where $\breve{A}$ is the restricted companion matrix of the reduced solution

$$\breve{A} = \begin{bmatrix} \tilde{\Theta}_1 & \tilde{\Theta}_2 & \cdots & \tilde{\Theta}_k \\ I_n & 0_{n\times n} & \cdots & 0_{n\times n} \\ \vdots & \ddots & \ddots & \vdots \\ 0_{n\times n} & 0_{n\times n} & I_n & 0_{n\times n} \end{bmatrix}$$

hence the CER become

$$\begin{cases} (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_1 - \Gamma_f\left(\tilde{\Theta}_1^2 + \tilde{\Theta}_2\right) - (\Gamma_b + R\Gamma_0) = 0_{n\times n} \\ (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_2 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_2 + \tilde{\Theta}_3\right) - (\Phi_2 - R\Gamma_b) = 0_{n\times n} \\ (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_3 - \Gamma_f\left(\tilde{\Theta}_1\tilde{\Theta}_3 + \tilde{\Theta}_4\right) - (\Phi_3 - R\Phi_2) = 0_{n\times n} \\ \vdots \\ (\Gamma_0 + R\Gamma_f)\,\tilde{\Theta}_k - \Gamma_f\tilde{\Theta}_1\tilde{\Theta}_k + R\Phi_{k-1} = 0_{n\times n} \end{cases}$$

■

8.4 Appendix D

Non-stationary case. In this Appendix we extend the analysis of the QR-NK-DSGE model to the situation in which Assumption 2 is replaced with:

Assumption 2' The characteristic equation $\det[\Theta(s)] = 0$ has $p - r$ roots equal to $s = 1$, $0 < r < p$ and the remaining roots are such that $\mid s \mid > 1$.

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 constants does not ensure stationarity if the persistence of the time series is governed by a unit root, see Cogley (2001), Juselius and Franchi (2007), Gorodnichenko and Ng (2011), Dees et al. (2008) and Fukač and Pagan (2009). Moreover, treating mistakenly nonstationary as stationary processes may flaw standard inferential procedures, see Johansen (2006), Li (2007), Fanelli (2008) and Fanelli and Palomba (2010). As pointed in Fanelli (2009) this Section shows how the QR-NK-DSGE model can be transormed, under a set of restrictions and without loss of information, to account for the cointegration restrictions which approximate the observed time series, improving on Kapetanios et al. (2007) and Fukač and Pagan (2009).

Assumption 2' implies that Z_t generated in eq.(27) is integrated of order one (I(1)). The system can be represented in Vector Error Correction (VEC) form

$$\Delta Z_t = \alpha\beta' Z_{t-1} + \Xi W_{t-1} + \epsilon_t \ , \ \epsilon_t \sim WN(0,\Sigma_\epsilon) \ , \ t = 1, ..., T \qquad (77)$$

where α and β are $p \times r$ matrices of full rank r respectively, and $\alpha\beta' = -(I_p - \Sigma_{j=1}^{k} A_j)$, $\Xi = [\Xi_1 : \Xi_2 : ... : \Xi_{k-1}]$, $\Xi_i = -\Sigma_{j=i+1}^{k} A_j$, $i = 1, ..., k-1$

and $W_{t-1} = (\Delta Z'_{t-1}, \Delta Z'_{t-2}, ..., \Delta Z'_{t-k+1})'$, see Johansen (1996). For $\beta = \beta^0$, where β^0 represents an identified version of the cointegration relations, the elements in $\beta_0' Z_t$ capture the stationary linear combinations of the variables in Z_t. Turning on the model presented in Section 1.2, if the output gap y_t is the only stationary variable in $Z_t = (y_t, \pi_t, i_t)'$, then $r = 1$ and $\beta_0 = (1, 0, 0)'$; if also the ex-post real interest rate is stationary, then $r = 2$ and $\beta_0 = (\beta_{01} : \beta_{02})$, where $\beta_{01} = (1, 0, 0)'$ and $\beta_{02} = (0, -1, 1)'$ and a single common stochastic trend drives the system. Once the cointegration rank r has been determined from the data (Cavaliere et al., 2012) and the hypothesis $\beta = \beta_0$ tested and not rejected, it is possible to define the $p \times 1$ (triangular) vector

$$Y_t := \begin{pmatrix} \beta_0' Z_t \\ \tau' \Delta Z_t \end{pmatrix} \quad \begin{matrix} r \times 1 \\ (p - r) \times 1 \end{matrix} \tag{78}$$

where τ 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 β_0 (Johansen, 1996). In the first example above, $Y_t = (y_t, \Delta\pi_t, \Delta i_t)'$ is obtained from eq.(78) 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. To represent the stationary reduced-form solution associated with the QR-NK-DSGE model, the structural equations can be reparameterized in terms of Y_t, i.e. such that only the error correction terms and the first difference of the variables are involved, see Fukač and Pagan (2009, Section 4.1, Strategy A). To do this, we re-write eq.(78) as

$$Y_t = P_{\beta_0,\Delta} Z_t \quad , \quad P_{\beta_0,\Delta} = \begin{bmatrix} \beta_0' \\ \tau'\Delta \end{bmatrix} \tag{79}$$

where $\Delta = (1 - L)$ and $P_{\beta_0,\Delta}$ is a $p \times p$ non-singular matrix; then we use $Z_t = P_{\beta_0,\Delta}^{-1} Y_t$ in system (28) and re-arrange the equations, obtaining

$$\begin{aligned} \Gamma_0^y Y_t &= \Gamma_f^y E_t Y_{t+1} + \Gamma_b^y Y_{t-1} + \sum_{j=2}^{k-1} \Phi_j Y_{t-j} \mathbb{I}_{\{k \geq 3\}} + \upsilon_t^y && (80) \\ \upsilon_t^y &= R\upsilon_{t-1}^y + \varepsilon_t^y, \quad \varepsilon_t^y \sim WN\left(0, \Sigma_{\varepsilon^y}\right) && (81) \end{aligned}$$

In system $(80) - (81)$, which can be regarded as the error-correcting counterpart of the QRE-NK-DSGE with I(1) cointegrated time series, the superscript 'y' remarks that other than being formulated in terms of Y_t, the parameters of the QR-NK-DSGE model accounts for all restrictions on β_0 and θ^* that ensure a balanced system. The estimation of the QR-NK-DSGE with I(1) cointegrated variables model can be carried out as follows. If the over-identifying restrictions characterizing β_0 are not rejected, the corresponding (Q)ML estimate $\widehat{\beta}_0$ can be used in place of β_0 in eq.(78) and treated as the 'true' value due to the super-consistency result (Johansen, 1996). The (Q)ML estimate of the vector of structural parameters θ_y^* can be obtained by applying the estimation algorithm described in Section 4.2 to system $(80) - (81)$. ■