

Giovanni Angelini, Luca Fanelli

Misspecification and Expectations
Correction in New Keynesian DSGE
Models

Quaderni di Dipartimento

Serie Ricerche 2015, n. 1

ISSN 1973-9346



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche “Paolo Fortunati”

Misspecification and Expectations Correction in New Keynesian DSGE Models*

Giovanni Angelini, Luca Fanelli

Department of Statistical Sciences & School of Economics, Management and Statistics
University of Bologna

February 2015

Abstract

This paper focuses on the dynamic misspecification that characterizes the class of small-scale New-Keynesian models currently used in monetary and business cycle analysis, and provides a ‘natural’ remedy for the typical difficulties these models have in accounting for the rich contemporaneous and dynamic correlation structure of the data, generally faced with *ad hoc* shock specifications. We suggest using the ‘best fitting’ statistical model for the data as a device through which it is possible to adapt the econometric specification of the New-Keynesian model, such that the risk of omitting important propagation mechanisms is kept under control. The statistical model may feature an autocorrelation structure that is more involved than the autocorrelation structure implied by the structural model’s reduced form solution under rational expectations, and it is treated as the actual agents’ expectations generating mechanism. A pseudo-structural form is built from the baseline system of Euler equations by forcing the state vector of the system to have the same dimension as the state vector characterizing the statistical model. Other than the structural parameters, the pseudo-structural form features a set of estimable parameters that correct the mismatch between rational and agents’ expectations. By construction, the pseudo-structural form gives rise to a set of cross-equation restrictions that do not penalize the autocorrelation structure found in the data. Standard estimation and evaluation methods can be used. We provide an empirical illustration based on U.S. quarterly data and a small-scale monetary New Keynesian model.

Keywords: Dynamic stochastic general equilibrium model, Expectations, Kalman filter, New Keynesian models, State space model.

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

*Paper presented at the ‘34th International Symposium on Forecasting’, Rotterdam, June 2014, and the ‘Sixth Italian Congress of Econometrics and Empirical Economics’, Salerno, January 2015. We wish to thank seminar participants for helpful comments and discussion. We are solely responsible for all remaining errors.

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 and inflation. It is well recognized, however, that these models capture only stylized features of the business cycle and the monetary policy stance, and display a limited time series performance. Indeed, despite possessing attractive theoretical properties, such as the capability of featuring potential structural sources of endogenous persistence, such as external habit persistence, implicit indexation, adjustment costs of investment, etc., NK-DSGE models are misspecified in several dimensions (Henry and Pagan, 2004; An and Schorfheide, 2007) and are typically treated as restricted, but parametrically incomplete representations of the actual data by econometricians (Diebold *et al.* 1998; An and Schorfheide, 2007; Canova and Ferroni, 2012). Although NK-DSGE models cannot provide a complete description of the business cycle, they are widely used to evaluate macroeconomic scenarios and to predict economic activity; therefore, assessing the correspondence between what these models imply and what the data tell us is a crucial step in the process of analysing policy options and their effects.

One important source of misspecification can be ascribed to the difficulties NK-DSGE models display in generating sufficient endogenous persistence to match the persistence observed in quarterly data. NK-DSGE models are built upon the rational expectations (RE) paradigm. Under RE, agents are assumed to know the data generating process (except the structural parameters) and form their expectations consistently. Two types of restrictions arise on the model's reduced form solution: (i) parametric nonlinear cross-equation restrictions (CER) that map the structural to the reduced form parameters; (ii) constraints on the lag order and correlation structure of the variables. The restrictions in (i) are the Hansen and Sargent's (1980, 1981) traditional 'metric' for the evaluation of models based on forward-looking behaviour and RE, see also Hansen (2014). Instead, the restrictions in (ii) are 'implicit', and very often, practitioners are not aware of their role and importance in the empirical evaluation of NK-DSGE models.

The unique stable solution associated with NK-DSGE models can be represented as a state space model, possibly transformed in minimal form (Komunjer and Ng, 2011), or as finite-order vector autoregressive (VAR) systems, in the special case in which all endogenous variables are observed. These solutions generally involve one (two) lag(s) of the endogenous variables, giving rise to what we call throughout the paper an 'omitted dynamics' issue. By this term, we denote the situation that occurs when the constraints in (ii) conflict with the propagation mechanisms one detects from the data using a statistical model that does not embody the

parametric constraints stemming from the theory. Testing the validity of the NK-DSGE model through the CER when the restrictions in (ii) conflict with the actual autocorrelation structure of the data might distort the evaluation process.

What should investigators do? The natural and obvious fix in these cases would require the estimation of a theoretically micro-founded model with less restrictive dynamics than the original New Keynesian model. An excellent example is provided in, e.g., Lubik and Schorfheide (2004), Section 5.D, who estimate a dynamically less restrictive version of their NK-DSGE model as a robustness check. They introduce a consumption Euler equation which features habit formation that generalizes the previously specified purely forward-looking consumption equation, and an ‘hybrid’ Phillips curve, as opposed to its purely forward-looking version. Examples like this, nevertheless, are rare, because it is not always practical to microfound all propagation mechanisms that characterize quarterly (or monthly) time series. What do practitioners typically do? They generally endow the shocks of the model with more elaborate and persistent time series models like, e.g., AR or ARMA-type processes (Smets and Wouters, 2007; Cúrdia and Reis, 2010), without (apparently) changing the specification of their structural equations.

The aim of this paper is to formalize a procedure that relaxes the extent of the restrictions in (ii) by relaxing the tightness of the RE paradigm. More specifically, we pursue the idea that only the CER in (i) should be considered and tested to evaluate the model, while the restrictions in (ii) should be neutralized when clearly at odds with the data. Our solution to the omitted dynamics issue is based on the idea of using a ‘best fitting’ state space model for the data as the actual agents’ expectations generating mechanism, without the need to resort to the adaptive learning framework (Evans and Honkapohja, 1999; 2001; 2003; Sargent, 1999; Branch and Evans, 2006; Milani, 2007). This leads to the definition of a pseudo-structural model that combines the structural information subsumed by the NK-DSGE model with features of the data captured by the statistical model. The pseudo-structural form is specified by augmenting the original system of Euler equations with additional lags of the variables, such that the gap between the dimension of the state vector under RE and the dimension of the state vector of the statistical model is filled up. By construction, the unique stable solution associated with the pseudo-structural model has the same time series representation as the statistical model, and it features a component that serves as the expectations correction term. By construction, the CER implied by the pseudo-structural model does not lead to restrictions in (ii) which are at odds with the agents’ forecasts. For ease of reference, we call the so-built pseudo-structural form the ‘NK-DSGE model under Quasi-RE (QRE)’.¹

¹We borrow the name Quasi Rational Expectations from Nerlove *et al.* (1979), Nelson and Blessler (1992), Nerlove and Fornari (1998) and Holt and McKenzie (2003), who introduced this concept in different fields of research. Strictly speaking, QRE would require replacing expectational variables in the structural equations with

Frequentist and Bayesian estimation and evaluation methods can easily be applied. The NK-DSGE model under RE is nested within the pseudo-structural model; hence, likelihood-ratio tests can be used to select the two specifications. More generally, information criteria or any other evaluation method can be exploited to select the specification that is more supported by the data. We propose an empirical illustration based on U.S. quarterly data, where we use the monetary business cycle model discussed in Benati and Surico (2009) as the reference structural model. We compare and evaluate the results obtained under RE and with our approach.

It is worth stressing that we do not propose the active use of a statistical model to rectify the specification of the NK-DSGE model as an end in itself. Rather, we see our approach as providing a useful specification check for NK-DSGE models, allowing a researcher to robustify inferences against one important dimension about the misspecification of the model, while capturing some important ‘stylized facts’. In this respect, our approach shares the viewpoint also adopted in Franchi and Juselius (2007) and, to some extent, in Consolo *et al.* (2009).²

Our paper has several connections with the existing literature. As already mentioned, one way to address the poor time series performance of structural forward-looking models has been advocated by Cúrdia and Reis (2010), who suggest augmenting the overall dynamics of macro business cycle models by allowing for disturbances that have a rich contemporaneous and dynamic correlation structure. In practice, they suggest replacing the usual unsatisfactory autoregressive specification of order one (AR(1)) of the model’s disturbances with more general AR or ARMA-type processes, allowing for possible cross-equation dependence, so as to maximize the best time series performance of the model. Similarly, Smets and Wouters (2007) specify ARMA(1,1)-type processes for the price mark-up and the wage mark-up disturbances in their medium-sized estimated DSGE model, observing that for these shocks the inclusion of the moving average terms is designed to capture the high-frequency fluctuations in inflation and wages. Like Cúrdia and Reis (2010), we let the data speak freely about the dimension of the dynamic misspecification of the system, but unlike Cúrdia and Reis (2010), our starting point is a statistical model for the data which is anchored to the theoretical model to make expectations

their values calculated from the ‘best fitting’ statistical model for them. The concept should be properly adapted in the context of dynamic stochastic general equilibrium models, see e.g., Fanelli (2009) for an early example.

²Broadly speaking, the best statistical model for the data might potentially (but not necessarily) also exploit information ‘external’ to the structural model, possibly derived from large datasets, think e.g., about factor models. For instance, Beyer *et al.* (2008) propose to combine factor analysis for information extraction from large data sets and generalized method of moments to estimate the parameters of systems of forward-looking equations. In principle, factor-augmented VAR models, as in Consolo *et al.* (2009), might be used as the agents’ expectations generating system. In this paper, we stick to the concept of model-consistent expectations; hence, it is assumed that the agents exploit only the information ‘internally’ recoverable from the structural model.

consistent with the agents' forecasts.

The most common and known alternative to RE is the adaptive learning hypothesis, see Evans and Honkapohja (1999; 2001; 2003), Sargent (1999), Branch and Evans (2006) and Milani (2007) for details. Under adaptive learning, agents are assumed to form and update their beliefs by using forecast models with time-varying coefficients and recursive updating rules. The postulated agents' forecasting model, or perceived law of motion, is typically (albeit not necessarily) the reduced form solution of the system under RE. Although the adaptive learning hypothesis can induce more persistence in the data (Branch and Evans 2006; Milani 2007; Chevillon *et al.* 2010), and it permits a substantial statistical relaxation of the strength of the CER (Fanelli 2008; Fanelli and Palomba 2011), a typical learning model focuses on the dynamic interaction between beliefs and observed data, and is not designed to solve the misspecification issue with which we are concerned in this paper.³ Moreover, as will be shown throughout the paper, our suggested approach does not deviate from the concept of model-consistent expectations. 'Consistency', however, is refers to the 'best-fitting' statistical model for the data.

Recently, Cole and Milani (2014) have investigated the ability of popular New Keynesian models to match the data in terms of their interaction between macroeconomic variables and their corresponding expectations. First, they report the failure of NK-DSGE models under RE to account for the dynamic interaction between macroeconomic expectations and macroeconomic realizations. Second, they observe that alternative models of expectations formation, including, e.g., extrapolative and heterogeneous expectations, can reconcile the NK-DSGE models with the data. Our approach represents another contribution towards the idea of reconciling the time series performance of NK-DSGE models with the data.

We also have some points in common with the DSGE-VAR approach of Del Negro *et al.* (2007). The DSGE-VAR approach is driven by the idea of assessing how far/close a dynamic macro model based on RE is from the data. Del Negro *et al.* (2007) propose a Bayesian evaluation method: They use a VAR system for the observed variables as the statistical model for the data, and centre the prior distribution for the VAR parameters on the CER implied by the structural model. The dispersion of these priors from the CER is governed by a scalar (hyper)parameter: Small values of such a (hyper)parameter indicate that the VAR is far from the theoretical model, while large values of this (hyper)parameter indicate that the theoretical model is supported by the data. In our setup, the statistical model that describes the data is either a VAR system or a state space model, depending on whether one can observe/proxy all endogenous variables or not. The statistical model determines the dynamic structure of the

³Cho and Kasa (2014) have recently proposed a model validation approach to learning, where agents operating in a self-referential environment are aware of potential model misspecification and try to detect it in real-time, using econometric specification tests.

pseudo-structural model that is confronted with the data. Like Del Negro *et al.* (2007), we are motivated by the idea of relaxing the tightness of the restrictions implied by the RE hypothesis, without renouncing the concept of model-consistent expectations.

The rest of the paper is organized as follows. In Section 2 we present our main idea through a simple uni-equational example. In Section 3 we introduce our prototype structural NK-DSGE model and discuss the omitted dynamics issue that arises under RE, and in Section 4 we present our approach. In Section 5 we estimate a NK-DSGE model for the U.S. economy using quarterly data: In Sub-section 5.1 we discuss the reference structural model, in Sub-section 5.2 we deal with ‘best fitting’ statistical model for the data and, finally, in Sub-sections 5.3 and 5.4 we address the frequentist and Bayesian estimation and evaluation results. Section 6 contains some concluding remarks.

2 Background

Consider a simple economy described by the uni-equational linear RE model

$$Z_t = \gamma_f E_t Z_{t+1} + \gamma_b Z_{t-1} + \omega_t, \quad \omega_t \sim \text{WN}(0, 1), \quad t = 1, \dots, T. \quad (1)$$

Z_t is an observable scalar generated by a covariance stationary process, Z_0 is given, $E_t Z_{t+1} := E(Z_{t+1} | \mathcal{F}_t)$ is the expectation operator conditional on the information set \mathcal{F}_t , and ω_t is a scalar white noise process with variance 1, called structural (or fundamental) disturbance (or structural shock). We call the model in Eq. (1) ‘structural model’. The structural parameters are $\gamma_f > 0$, $\gamma_b > 0$, and are collected in the vector $\theta := (\gamma_f, \gamma_b)'$.

Assuming that $\gamma_f + \gamma_b < 1$, the unique stable RE solution to the model in Eq. (1) is given by the autoregressive model of order one (AR(1)):

$$Z_t = \tilde{\phi} Z_{t-1} + \tilde{\psi} \omega_t, \quad t = 1, \dots, T \quad (2)$$

where $\tilde{\phi} = \phi(\theta)$ and $\tilde{\psi} = \psi(\theta)$ are reduced form parameters that depend nonlinearly on θ . A ‘tilde’ over ϕ and ψ is used to stress the fact that these parameters are forced to depend on θ under RE. In particular, $\tilde{\phi}$ is the real stable root (i.e. $\tilde{\phi} \in (-1, 1)$) of the second-order equation $\gamma_f \phi^2 - \phi + \gamma_b = 0$, and $\tilde{\psi} = (1 - \gamma_f \tilde{\phi})^{-1}$.

Under RE, the data generating process belongs to the class of models described by Eq. (2). Consistent estimates of θ can be obtained from the autoregressive parameter $\tilde{\phi}$ and the variance σ_ε^2 of $\varepsilon_t := \tilde{\psi} \omega_t$, by imposing the CER: $\phi = \tilde{\phi}$, $\sigma_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2$, where $\tilde{\phi}$ is the real stable solution to $\gamma_f \phi^2 - \phi + \gamma_b = 0$ and $\tilde{\sigma}_\varepsilon^2 = (1 - \gamma_f \tilde{\phi})^{-2}$. Moreover, the autocorrelation structure of the time series Z_1, Z_2, \dots, Z_T should conform to that of AR(1)-type processes.

Assume now that based on his/her specification analysis, the econometrician believes the data generating process belongs to the class of covariance stationary AR(2) processes of the form

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t \quad , \quad \varepsilon_t \sim \text{WN}(0, \sigma_\varepsilon^2) \quad , \quad t = 1, \dots, T, \quad (3)$$

where the autoregressive coefficient associated with the second lag, ϕ_2 , is such that $\phi_2 \neq 0$. We call the model in Eq. (3) the statistical model for the data. The parameters of the statistical model are $\tau := (\phi_1, \phi_2, \sigma_\varepsilon^2)'$. Compared to the reduced form solution in Eq. (2), the AR(2) model in Eq. (3) involves an additional lag of the state variable Z_t . For the econometrician, the best forecast of Z_{t+1} conditional on the information set available at time t will be $E(Z_{t+1} | \mathcal{F}_t) = \phi_1 Z_t + \phi_2 Z_{t-1}$ and not $E(Z_{t+1} | \mathcal{F}_t) = \phi_1 Z_t$ as predicted by the structural model under RE.

Since the model in Eq. (2) is nested in Eq. (3), the AR(2) model might be interpreted as the reduced form solution associated with the structural model in Eq. (1) if the following set of restrictions hold:

$$\begin{aligned} \text{res-I:} \quad & \phi_1 = \tilde{\phi}, \text{ where } \tilde{\phi} \text{ is the stable root of } \gamma_f \phi^2 - \phi + \gamma_b = 0 \\ & \sigma_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2, \text{ where } \tilde{\sigma}_\varepsilon^2 = (1 - \gamma_f \tilde{\phi})^{-2} \end{aligned} \quad (4)$$

$$\text{res-II:} \quad \phi_2 = \tilde{\phi}_2 = 0.$$

In principle, the structural parameters θ might be estimated consistently from the model in Eq. (3) by imposing the restrictions in Eq. (4). It is clear, however, that the restrictions res-II in Eq. (4) conflict with the econometrician's finding that $\phi_2 \neq 0$. If the data generating process belongs to the class of models in Eq. (3) with $\phi_2 \neq 0$, the estimator of θ recovered from model Eq. (2) imposing the CER in Eq. (4) will be distorted because of the omission of a relevant regressor.

The natural fix to this shortcoming should be the re-specification of a theory-based structural model implying a time series representation for Z_t featuring Z_{t-2} , other than Z_{t-1} . Yet only seldom is that feasible. We discuss two solutions to the 'omitted dynamics' issue: First, the 'conventional' approach, and subsequently our solution.

Conventional approach

The 'conventional' approach works by endowing the structural model in Eq. (1) with an AR(1) process for the shocks, now denoted with ω_t^* , i.e.

$$\begin{aligned} Z_t &= \gamma_f E_t Z_{t+1} + \gamma_b Z_{t-1} + \omega_t^* \quad , \quad t = 1, \dots, T \\ \omega_t^* &= \rho \omega_{t-1}^* + v_t \quad , \quad |\rho| < 1 \quad , \quad v_t \sim \text{WN}(0, 1 - \rho^2). \end{aligned} \quad (5)$$

In this specification, ρ is an autoregressive parameter and v_t is the structural shock (which is normalized such that the variance of ω_t^* is still equal to 1). The autoregressive equation for ω_t^*

and the associated autoregressive parameter, ρ , are not generally derived from first principles, but from the practical purpose of improving the statistical fit of the model. Apparently, the theoretical structural model in Eq. (1) has not been changed. Actually, by exploiting the autoregressive structure of ω_t^* and using simple algebra, we obtain:

$$Z_t = \frac{\gamma_f}{1 + \rho\gamma_f} E_t Z_{t+1} + \frac{\gamma_b + \rho}{1 + \rho\gamma_f} Z_{t-1} - \frac{\rho\gamma_b}{1 + \rho\gamma_f} Z_{t-2} + \frac{1}{1 + \rho\gamma_f} v_t^* \quad , \quad t = 1, \dots, T \quad (6)$$

where $v_t^* := \rho\gamma_f\eta_t + v_t$, and $\eta_t := Z_t - E_{t-1}Z_t$ is a martingale difference sequence forecast error ($E_{t-1}\eta_t = 0$). The representation in Eq. (6) recalls a well known fact from textbook econometrics: Autoregressive disturbances amount to additional lagged regressors of the endogenous variable. The unique stable RE solution associated with Eq. (6) is given by the AR(2) process in Eq. (3), with τ subject to the following set of CER: $\phi_1 = \tilde{\phi}_1$, $\phi_2 = \tilde{\phi}_2$, $\sigma_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2$, where

$$\begin{aligned} (1 - \gamma_f\tilde{\phi}_1 + \rho\gamma_f)\tilde{\phi}_1 &= (\gamma_f\tilde{\phi}_2 + \gamma_b + \rho) \\ (1 - \gamma_f\tilde{\phi}_1 + \rho\gamma_f)\tilde{\phi}_2 &= -\rho\gamma_b \\ \tilde{\sigma}_\varepsilon^2 &= \left(\frac{1-\rho}{1-\gamma_f\tilde{\phi}_1} \right)^2. \end{aligned} \quad (7)$$

It can be noticed that $\rho \neq 0$ implies $\phi_2 \neq 0$. Instead, if $\rho = 0$, the restrictions above collapse to those in Eq. (4).

Suggested approach

Our suggested approach is based on a slight change of perspective. We assume that the AR(2) model in Eq. (3) is the agents' forecast model, and introduce a 'pseudo-structural' form whose reduced form solution has a time series representation consistent with Eq. (3). The pseudo-structural form is in this case given by

$$Z_t = \gamma_f E_t Z_{t+1} + \gamma_b Z_{t-1} + \zeta Z_{t-2} + \omega_t^{**}, \quad \omega_t^{**} \sim \text{WN}(0, 1), \quad t = 1, \dots, T, \quad (8)$$

and is obtained from the structural model in Eq. (1) by adding the term ζZ_{t-2} to the baseline equation. The disturbance ω_t^{**} is still a white noise term with variance 1.

The crucial question here is: How do we interpret the ζZ_{t-2} term in Eq. (1)? In principle, ζZ_{t-2} might be interpreted as a term capturing propagation mechanisms that are not directly explained by the theoretical model because of the omission of adjustment costs, information delays, time-to-build effects, etc. These effects, however, should be modelled endogenously in the structural specification, if any. In our setup, ζZ_{t-2} plays the role of an 'expectations correction' term in a sense that will be qualified below. The auxiliary parameter ζ is not constrained to lie in the (-1,1) interval; the only implicit requirement we give to ζ is that if for a given $\theta = \check{\theta}$ the solution to the model in Eq. (1) is unique and stable, the solution to the model in Eq. (8) also must be unique and stable.

The specification in Eq. (8) should be no more disturbing than that in Eq. (6) obtained by adding the autoregressive disturbance to the structural theoretical model. The unique stable solution associated with the model in Eq. (8), if it exists, is given by the AR(2) process in Eq. (3) with parameters τ subject to the following set of CER: $\phi_1 = \tilde{\phi}_1$, $\phi_2 = \tilde{\phi}_2$, $\sigma_\varepsilon^2 = \tilde{\sigma}_\varepsilon^2$, where

$$\begin{aligned} (1 - \gamma_f \tilde{\phi}_1) \tilde{\phi}_1 &= (\gamma_f \tilde{\phi}_2 + \gamma_b) \\ (1 - \gamma_f \tilde{\phi}_1) \tilde{\phi}_2 &= \zeta \\ \tilde{\sigma}_\varepsilon^2 &= (1 - \gamma_f \tilde{\phi}_1)^{-2}. \end{aligned} \tag{9}$$

It can be noticed that, in this case, $\zeta \neq 0$ also implies $\phi_2 \neq 0$. Instead, if $\zeta = 0$, the restrictions above collapse to those in Eq. (4).

Although the likelihoods associated with the AR(2) model under the restrictions in Eq. (7) and in Eq. (9) may be numerically different (recall that ρ lies in the (-1,1) interval, while ζ is subject to another requirement), at first glance, the two approaches seem to be equivalent. Yet they are conceptually different. With the ‘conventional’ approach, the practitioner does not need to specify any statistical model for the data. He/she will specify a time series process for the disturbance ω_t with the aim of improving the overall empirical fit of the model. Our approach is instead based on the idea of treating the statistical model for the data like the actual agents’ expectations generating system. The term ζZ_{t-2} in Eq. (8) plays the role of an expectations correction term that fills up the mismatch between rational and agents’ expectations. The term ζZ_{t-2} guarantees that the unique stable solution associated with the pseudo-structural form has the same time series representation as the agents’ statistical model for the data, and that the difference between the two models is only due to the CER.

3 The NK-DSGE model under Rational Expectations and the omitted dynamics issue

Let $Z_t := (Z_{1,t}, Z_{2,t}, \dots, Z_{n,t})'$ be a $n \times 1$ vector of endogenous variables and assume that after log-linearization, the structural form of the NK-DSGE model can be represented in the form

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + C + \eta_t \quad , \quad t = 1, \dots, T \tag{10}$$

where, $\Gamma_i := \Gamma_i(\theta)$, $i = 0, f, b$ are $n \times n$ matrices whose elements depend on the vector of structural parameters θ , $C := C(\theta)$ is a $n \times 1$ constant which can be non-zero when it is intended to capture steady state values of some variables of the system, and η_t is a $n \times 1$ vector of disturbances 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 , $E_t Z_{t+1} := E(Z_{t+1} | \mathcal{F}_t)$. Without any loss of generality, Γ_0 is assumed non-singular. When a direct link between the process generating η_t and a set of 'forcing variables' is not provided by the theory, a typical completion of system (10) is obtained through the autoregressive specification

$$\eta_t = R\eta_{t-1} + \omega_t \quad , \quad \omega_t \sim \text{WN}(0, \Sigma_\omega) \quad (11)$$

where R is a $p \times p$ diagonal stable matrix (i.e. with its eigenvalues inside the unit disk), and ω_t is a white noise term with covariance matrix Σ_ω that can be diagonal or non-diagonal. The true value of θ , θ_0 , is assumed to be an interior point of the parameter space Θ .

The multivariate linear RE model in Eq.s (10)-(11) nests a large class of small-scale linearized NK-DSGE models used in monetary policy analysis. There exists many solution methods available in the literature by which a reduced form solution of system (10)-(11) can be computed under RE. A solution of system (10)-(11) is any stochastic process $\{Z_t^*\}_{t=0}^\infty$ such that, for $\theta \in \Theta$, $E_t Z_{t+1}^* = E(Z_{t+1}^* | \mathcal{F}_t)$ exists and if Z_t^* is substituted for Z_t into the structural equations, the model is verified for each t , for fixed initial conditions. A reduced form solution is a member of the solution set whose time series representation is such that Z_t can be expressed as a function of ω_t , lags of Z_t and ω_t and, possibly, other arbitrary martingale difference sequences (MDS) with respect to \mathcal{F}_t , independent of ω_t , called 'sunspot shocks', see Fanelli (2012) and Castelnuovo and Fanelli (2015).

Assuming that θ_0 lies in the determinacy region of Θ , the unique stable reduced form solution associated with system (10)-(11) can be represented in the form (see Binder and Pesaran, 1995; Uhlig, 1999; Klein, 2000)

$$\begin{pmatrix} Z_t - \tilde{v} \\ Z_{t-1} - \tilde{v} \\ \vdots \\ x_t \end{pmatrix} = \begin{pmatrix} \tilde{\Phi}_1 & \tilde{\Phi}_2 \\ I_n & 0_{n \times n} \end{pmatrix} \begin{pmatrix} Z_{t-1} - \tilde{v} \\ Z_{t-2} - \tilde{v} \\ \vdots \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \tilde{\Psi} \\ 0_{n \times n} \end{pmatrix} \omega_t \quad (12)$$

where $\tilde{v} := (I_n - \tilde{\Phi}_1 - \tilde{\Phi}_2)^{-1} \tilde{\mu}$, and we use 'tildes' over the matrices of parameters to remark the fact that $\tilde{\Phi}_1, \tilde{\Phi}_2, \tilde{\Psi}$ and $\tilde{\mu}$ depend on θ through the set of CER:

$$\begin{aligned} (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_1 - \Gamma_f \tilde{\Phi}_2 + \Gamma_{b,1} &= 0_{n \times n} \\ (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_2 - \Gamma_{b,2} &= 0_{n \times n} \end{aligned} \quad (13)$$

$$\begin{aligned} \tilde{\mu} - (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1 - \Gamma_f) C &= 0_{n \times 1} \\ \tilde{\Sigma}_\varepsilon = \tilde{\Psi} \Sigma_\omega \tilde{\Psi}' \quad , \quad \tilde{\Psi} &= \left(\Gamma_0 - \Gamma_f \tilde{\Phi}_1 \right)^{-1}. \end{aligned} \quad (14)$$

In the expressions in Eq.s (13)-(14), $\Gamma_0^R = (\Gamma_0 + R\Gamma_f)$, $\Gamma_{b,1} = (\Gamma_b + R\Gamma_0)$, $\Gamma_{b,2} = -R\Gamma_b$, and $\tilde{\Sigma}_\varepsilon$ is the covariance matrix of the reduced form disturbance $\varepsilon_t = \tilde{\Psi}\omega_t$, see Bårdsen and Fanelli (2015) and Castelnovo and Fanelli (2015). A convenient way to summarize the equilibrium in Eq.s (12)-(14) is to refer to the representation

$$\begin{matrix} x_t \\ 2n \times 1 \end{matrix} = \begin{matrix} A(\theta) & G(\theta) \\ 2n \times 2n & 2n \times n \end{matrix} \begin{matrix} x_{t-1} \\ \omega_t \end{matrix} + \begin{matrix} 2n \times 1 \\ n \times 1 \end{matrix} \quad (15)$$

Let $y_t := (y_{1,t}, y_{2,t}, \dots, y_{p,t})'$ be the $p \times 1$ vector of observable variables. When all variables in Z_t are observed, $y_t = Z_t$, and the transition system in Eq. (15) along with the measurement system: $y_t = Hx_t$, $H := (I_n : 0_{n \times n})$, give rise to a VAR representation for Z_t in which the VAR coefficients depend on θ through the CER in Eq.s (13)-(14). In general, when not all variables in Z_t are observed, the measurement system will take the form

$$y_t = Hx_t + Vv_t \quad (16)$$

where H is a $p \times 2n$ matrix, v_t a $b \times 1$ vector ($b \leq p$) of measurement errors with covariance matrix Σ_v , and V is a $p \times b$ selection matrix. Let $u_t := (\omega_t', v_t')'$ be the $n + b$ complete vector of innovations. By substituting Eq. (15) into Eq. (16) and using some algebra, one obtains the the so-called ABCD representation

$$\begin{matrix} x_t \\ 2n \times 1 \end{matrix} = \begin{matrix} A(\theta) & B(\theta) \\ 2n \times 2n & 2n \times (n+b) \end{matrix} \begin{matrix} x_{t-1} \\ u_t \end{matrix} + \begin{matrix} 2n \times 1 \\ (n+b) \times 1 \end{matrix} \quad (17)$$

$$\begin{matrix} y_t \\ p \times 1 \end{matrix} = \begin{matrix} C(\theta) & D(\theta) \\ p \times 2n & 2n \times (n+b) \end{matrix} \begin{matrix} x_{t-1} \\ u_t \end{matrix} + \begin{matrix} p \times 1 \\ (n+b) \times 1 \end{matrix}$$

where $B(\theta) := (G(\theta) : 0_{2n \times b})$, $C(\theta) := HA(\theta)$ and $D(\theta) := (HG(\theta) : V)$.⁴

The state space system (17) summarizes the determinate (unique and stable) equilibrium associated with the NK-DSGE model under RE. (It collapses to a VAR for Z_t when $y_t = Z_t$). Provided θ is locally identifiable, the state space model defined by Eq.s (15)-(16) can be taken to the data using different estimation methods, see e.g. Ruge-Murcia (2007). 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 dynamics issue we are concerned with in this paper, assume that

⁴We refer to Fernández-Villaverde et al. (2007), Ravenna (2007), Franchi and Vidotto (2013) and Franchi and Paruolo (2014) for a detailed analysis of the cases in which y_t can be given a fundamental and finite-order VAR representation when the $D = D(\theta)$ matrix in Eq. (17) is square. More generally, the state space model in Eq. (17) will give rise to VARMA-type representations for y_t , see e.g. Hannan and Deistler (1988).

the ‘best fitting’ model for the data is represented by the state space system

$$\begin{pmatrix} Z_t - v \\ Z_{t-1} - v \\ \vdots \\ Z_{t-k+1} - v \\ x_t^* \end{pmatrix} = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{k-1} & \Phi_k \\ I_n & 0_{n \times n} & \cdots & 0_{n \times n} & 0_{n \times n} \\ & \ddots & & \vdots & \vdots \\ 0_{n \times n} & 0_{n \times n} & \cdots & I_n & 0_{n \times n} \end{pmatrix} \begin{pmatrix} Z_{t-1} - v \\ Z_{t-2} - v \\ \vdots \\ Z_{t-k} - v \\ x_{t-1}^* \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0_{n \times n} \\ \vdots \\ 0_{n \times n} \end{pmatrix} \quad (18)$$

$$y_t = H^* x_t^* + V^* v_t \quad (19)$$

where $\Phi_1, \Phi_2, \dots, \Phi_k, \mu$ and $\Sigma_\varepsilon := E(\varepsilon_t \varepsilon_t')$ are matrices of coefficients in which no theoretical restriction is placed, $v := (I_n - \Phi_1 - \Phi_2 - \dots - \Phi_k)^{-1} \mu$, and H^* and V^* are selection matrices of suitable dimensions. It is assumed that $\Phi_k \neq 0_{n \times n}$. Collecting the parameters of system (18)-(19) in the vector $\tau := (\text{vec}(\Phi_1)', \dots, \text{vec}(\Phi_k)', \mu', \text{vech}(\Sigma_\varepsilon)', \text{vech}(\Sigma_v)')$ and using the vector of innovations $u_t := (\omega_t', v_t')$, also the statistical model can be summarized in ABCD form:

$$\begin{aligned} x_t^* &= A^*(\tau) x_{t-1}^* + B^*(\tau) u_t \\ y_t &= C^*(\tau) x_{t-1}^* + D^*(\tau) u_t \end{aligned} \quad (20)$$

that is assumed to be in ‘minimal form’, see Komunjer and Ng (2011).⁵ Again, we assume that τ is locally identifiable (Komunjer and Ng, 2011) and that the state space model in Eq. (20) collapses to a stationary VAR for Z_t when $y_t = Z_t$.

The simple comparison of systems (17) and (20) reveals that the dimension of the state vector in Eq. (20) will be generally larger than the dimension of the state vector in system (17), i.e. $\dim(x_t^*) \geq \dim(x_t)$. The CER that the NK-DSGE model in Eq.s (10)-(11) places on τ can be represented in the form

$$(\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_1 - \Gamma_f \tilde{\Phi}_2 + \Gamma_{b,1} = 0_{n \times n} \quad (21)$$

$$(\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_2 - \Gamma_{b,2} = 0_{n \times n}$$

$$C - (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1 - \Gamma_f) \tilde{\mu} = 0_{n \times 1}$$

$$\tilde{\Sigma}_\varepsilon = \tilde{\Psi} \Sigma_\omega \tilde{\Psi}' \quad , \quad \tilde{\Psi} := (\Gamma_0 - \Gamma_f \tilde{\Phi}_1)^{-1} \quad (22)$$

$$\Phi_j = \check{\Phi}_j = 0_{n \times n} \quad , \quad j = 3, 4, \dots, k. \quad (23)$$

The restrictions in Eq.s (21)-(22) coincide with those in Eq.s (13)-(14), but now we have the additional set of $n^2(k-2)$ zero restrictions, summarized in Eq. (23), that force the dimension of the state vector x_t^* of the statistical model to match the dimension of x_t in the structural model.

⁵Minimality means that the model involves a minimum (non redundant) number of state variables. In practice, this conditions corresponds to ruling out common (cancelling) roots from VARMA-type systems.

While the CER in Eq.s (21)-(22) define a nonlinear mapping from θ to τ (res-I), say $\tau = g(\theta)$, where $g(\cdot)$ is a nonlinear continuous vector function, the zero restrictions in Eq. (23) imply $\dim(x_t^*) = \dim(x_t)$ (res-II). When in particular the data generating process belongs to the class of models defined by system (18)-(19) (or equivalently system (20)) and $k \geq 3$, $\dim(x_t^*) > \dim(x_t)$ and the CER in Eq.s (13)-(22) lead to the omitted dynamics issue.

4 The pseudo-structural form

Consider the NK-DSGE model in Eq.s (10)-(11) and the statistical model in Eq.s (18)-(19). We exploit the information stemming from the latter by introducing the following assumptions:

Assumption 1 [Expectations generating mechanism] Agents form their expectations using system (18)-(19).

Assumption 2 [Data generating process] The data generating process belongs to the class of models in Eq.s (18)-(19) with $\dim(x_t^*) = nk^{op}$ and $\Phi_{kop} \neq 0_{n \times n}$; the associated ABCD representation in Eq. (20) is in minimal form.

Assumption 3 [Stationarity] The matrix $A^*(\tau)$ is stable.

Assumption 4 [Parameters invariance] The parameters in τ does not vary over the period $t = 1, 2, \dots, T$.

Assumption 1 postulates that the agents form their expectations consistently with the statistical model for the data. Assumption 2 is a correct specification hypothesis. It maintains that the data generating process belongs to the specified statistical model, and that such a model involves the minimum number of state variables necessary to capture the propagation mechanisms at work in the data. We remark that Assumptions 1-2 are not in contrast with the concept of model-consistent expectations, as it will be clear below. Assumption 3 implies that the statistical model is asymptotically stable. The analysis can be easily extended to the case of unit roots and cointegration when $y_t = Z_t$, along the lines discussed in e.g. Fanelli (2009), Fukač and Pagan (2009) and Bårdsen and Fanelli (2015). Finally, Assumption 4 postulates that the parameters of the statistical model are time invariant.⁶

⁶At first glance, Assumption 4 is at odds with the logic of the adaptive learning hypothesis. Indeed, in the case of learning, the (population) parameters of the perceived law of motion - the agents' beliefs - are typically treated as time-varying coefficients which are updated recursively as new information become available. However, if in our case the 'best fitting' model for the data would be specified with drifting parameters τ_t , and its associated law of motion would be associated with the evolution of the agents' beliefs, our setup would line up with a the adaptive learning hypothesis.

Given the structural form in Eq.s (10)-(11) and the statistical model in Eq.s (18)-(19), we build a ‘pseudo-structural’ form associated with the NK-DSGE model, given by:

$$\begin{cases} \Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \left(\sum_{j=2}^{k^{op}-1} \Upsilon_j Z_{t-j} \right) \mathbb{I}_{\{k^{op} \geq 3\}} + C + \eta_t \\ \eta_t = R\eta_{t-1} + \omega_t \end{cases} \quad (24)$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function, and the $n \times n$ matrices Υ_j , $j = 2, \dots, k^{op} - 1$ contain, when $k^{op} \geq 3$, additional auxiliary parameters associated with $k^{op} - 2$ additional lags of Z_t . To keep the number of auxiliary parameters as small as possible, the matrices Υ_j s can be specified diagonal. When $k^{op} \leq 2$, the pseudo-structural form coincides with the ‘conventional’ NK-DSGE model in Eq.s (10)-(11).⁷ It turns out that the NK-DSGE model in Eq.s (10)-(11) is nested within the so-built pseudo-structural form.

It is tempting to interpret the term $\left(\sum_{j=2}^{k^{op}-1} \Upsilon_j Z_{t-j} \right) \mathbb{I}_{\{k^{op} \geq 3\}}$ in Eq. (24) as a component summarizing the effects of propagation mechanisms that albeit present in the data, are omitted by the baseline structural specification, such as length of real contracts, adjustment costs, delays in information flows, decision lags, etc., see e.g. Kozicki and Tinsley (1999), Rudebusch (2002*a*, 2002*b*) and Fuhrer and Rudebusch (2004) for example. However, if actually important, these effects should be micro-founded and incorporated directly in the structural specification. In our setup, the quantity $\left(\sum_{j=2}^{k^{op}-1} \Upsilon_j Z_{t-j} \right) \mathbb{I}_{\{k^{op} \geq 3\}}$ in Eq. (24) plays the role of an expectations correction term. It forces the reduced form solution associated with system (24) to be consistent with the postulated expectations generating system (Assumptions 1-2), as Proposition 1 below will clarify.

To fully understand the nature of system (24), we focus on its i -th Euler equation for $k^{op} \geq 3$, which is given by

$$Z_{i,t} = \gamma'_{i,0} Z_{i,t}^* + \gamma'_{i,f} E_t Z_{t+1} + \gamma'_{i,b} Z_{t-1} + \left(\sum_{j=2}^{k^{op}-1} \zeta'_{i,j} Z_{i,t-j} \right) + C_i + \eta_{i,t}$$

$$\eta_{i,t} = R_i \eta_{i,t-1} + \omega_{i,t} \quad , \quad i = 1, \dots, n.$$

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 , $\zeta'_{i,j}$ is the i -th diagonal element of Υ_j , $j = 1, \dots, k - 1$, C_i is the i -th element of C and, finally, $\eta_{i,t}$ and $\omega_{i,t}$ are the i -th elements of the vectors η_t and ω_t , respectively, where the autoregressive parameter $-1 < R_i < 1$ is the i -th diagonal component of R .

⁷In our setup the case $k^{op} = 1$ coincides with the situation where $R = 0_{n \times n}$ in Eq. (11).

Let ζ be the vector collecting the expectations correction parameters contained in the matrices Υ_j , $j = 2, \dots, k^{op}-1$, and $\theta^* = (\theta', \zeta')$ the vector containing all parameters associated with the pseudo-structural form in Eq. (24). The next assumption ensures that the uniqueness/stability properties of the NK-DSGE model are inherited by the associated pseudo-structural form.

Assumption 5 [Invariance to determinacy] Given the pseudo-structural form in Eq. (24), the expectations correction parameters in ζ are restricted such that for any $\theta = \check{\theta} \in \Theta$ for which a determinate solution for the NK-DSGE model in Eq.s (10)-(11) exists, a solution to system (24) exists and is determinate.

The next proposition derives the model-consistent reduced form solution and the CER implied by the pseudo-structural form in Eq. (24).

Proposition 1 [The CER under QRE] Under Assumptions 1-5, if a unique stable reduced form solution exists for the model in Eq. (24), it can be represented as in the form (18)-(19) with τ subject to the following set of CER: $\Phi_j = \check{\Phi}_j$, $j = 1, \dots, k$, $\mu = \check{\mu}$, $\Sigma_\varepsilon = \check{\Sigma}_\varepsilon$, where

$$\left\{ \begin{array}{l} (\Gamma_0^R - \Gamma_f \check{\Phi}_1) \check{\Phi}_1 - (\Gamma_f \check{\Phi}_2 + \Gamma_{b,1}) = 0_{n \times n} \\ (\Gamma_0^R - \Gamma_f \check{\Phi}_1) \check{\Phi}_2 - (\Gamma_f \check{\Phi}_3 + \Gamma_{b,2} + \Upsilon_2) = 0_{n \times n} \\ (\Gamma_0^R - \Gamma_f \check{\Phi}_1) \check{\Phi}_3 - (\Gamma_f \check{\Phi}_4 + \Upsilon_3 - R\Upsilon_2) = 0_{n \times n} \\ \vdots \\ (\Gamma_0^R - \Gamma_f \check{\Phi}_1) \check{\Phi}_{k^{op}} + R\Upsilon_{k^{op}-1} = 0_{n \times n} \\ (\Gamma_0^R - \Gamma_f \check{\Phi}_1 - \Gamma_f) \check{\mu} - (I_n - R)C = 0_{n \times 1} \\ \check{\Sigma}_\varepsilon - \check{\Psi} \Sigma_\omega \check{\Psi}' = 0_{n \times n} \quad , \quad \check{\Psi} := (\Gamma_0 - \Gamma_f \check{\Phi}_1)^{-1}. \end{array} \right. \quad (25)$$

Proof: See Appendix A.

The interesting feature of the restrictions in Eq. (25) is that there are no zero restrictions that reduce the length of the vector x_t^* in system (18)-(19).⁸

To distinguish our approach from the case in which expectations are computed directly from the reduced form solution associated with the structural model in Eq.s (10)-(11), hereafter we denote the pseudo-structural form under the restrictions of Proposition 1 as the ‘NK-DSGE

⁸A natural concern here is whether the CER in Eq. (25) allow to identify θ^* . A convenient way to summarize the CER of Proposition 1 is by the distance function $f(\tau, \theta^*) = 0_{a \times 1}$, where $f(\cdot, \cdot)$ is a nonlinear continuous vector differentiable function and $a = n^2 k^{op} + n + \frac{1}{2}n(n+1)$. By the implicit function theorem, the CER can be represented in explicit form $\tau = g(\theta^*)$, where $g(\cdot)$ is a nonlinear continuous differentiable vector function. Although an analytic expression for the function $g(\cdot)$ is not generally available, the Jacobian of the relationship can be computed with minor adaptations along the lines of Bårdsen and Fanelli (2015).

model under Quasi Rational Expectations (QRE)'. Such a name is motivated by the observation that in our setup the statistical model in Eq.s (18)-(19) is treated likewise the true agents' expectations generating mechanism (Assumption 1). In practise, our approach does not lead to any deviation from the concept of model-consistent expectations.

A frequentist econometric procedure for the NK-DSGE under QRE can be based on the following two steps:

- Step 1 Fit the statistical model in Eq.s (18)-(19) to the data, and use information criteria or likelihood-ratio tests to determine the dimension of the state vector, $\dim(x_t^*) = nk^{opt}$, that fits the data optimally. This can be done by estimating the state space model through maximum likelihood and Kalman filtering. The specific procedure we use to find a global maximum for τ 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 estimate $\hat{\tau}$. If it is found that $k^{opt} \leq 2$, the NK-DSGE model is estimated and evaluated in the 'conventional' way, i.e. under RE. If it is found that $k^{opt} \geq 3$, consider the next step;
- Step 2 Given $\dim(x_t^*) = nk^{opt}$, estimate $\theta^* = (\theta', \zeta')'$ from the statistical model in Eq.s (18)-(19) under a numerical approximation of the CER in Eq. (25), and then test the CER through a likelihood ratio test that compares the likelihood obtained in the previous step, $\log L_T(\hat{\tau})$, and the likelihood associated with θ^* , $\log L_T(\hat{\theta}^*)$, obtaining $LR_T^{CER} := -2(\log L_T(\hat{\theta}^*) - \log L_T(\hat{\tau}))$. The log-likelihood maximization is also achieved through Kalman filtering and Andreasen's (2010) algorithm.

If in Step 2 the CER are not rejected and standard regularity conditions hold (other than Assumptions 1-4), the estimator of θ^* (hence of θ) is consistent and asymptotically Gaussian, while LR_T^{CER} is asymptotically $\chi^2(d)$, with $d = \dim(\tau) - \dim(\theta^*)$. Moreover, the possible rejection of the CER can not be ascribed to the omitted dynamics issue.⁹

The Step 2 of the procedure can easily be adapted to the Bayesian approach. Given the statistical model built in Step 1, it is possible to specify a prior distribution for θ^* , $p(\theta^*)$, and then compute the posterior given the observations y_1, \dots, y_T , obtaining $p(\theta^* | y_1, \dots, y_T)$. This can be done by using e.g. the Random Walk Metropolis (RWM) algorithm along the lines of An

⁹Standard regularity conditions might not hold because of sample identification issues documented by e.g. Canova and Sala (2009). In these cases, it is possible to adapt the two steps above by using identification-robust methods along the lines of Dufour *et al.* (2013), Guerron-Quintana *et al.* (2013), Andrews and Mikusheva (2014) and Castelnuovo and Fanelli (2015).

and Schorfheide (2007). In our framework, it is ‘natural’ to specify priors for the expectations correction parameters ζ that are centered on zero, i.e. on the RE solution, such that the extent of the misspecification, if any, is determined by the data, see the next sections. Moreover, the selection between the NK-DSGE model under RE and QRE can be based on Bayesian information criteria or odds-ratios, etc. More details are provided in Sub-section 5.4.

5 Empirical analysis

In this section, we estimate and empirically evaluate a small-scale monetary NK-DSGE model on U.S. quarterly data, applying the QRE methodology discussed in the previous section. We also compare our approach with the ‘conventional’ RE case. In Sub-section 5.1 we introduce the reference structural model. In Sub-section 5.2 we describe the data and discuss the specification of the agents’ statistical model. In Sub-section 5.3 we estimate and evaluate the NK-DSGE model under QRE using a frequentist maximum likelihood approach, and in Sub-section 5.4 we repeat the same exercise using the Bayesian approach.

5.1 Structural model

Our reference NK-DSGE model is taken from Benati and Surico (2009) and is based on the following three equations:

$$\tilde{o}_t = \gamma E_t \tilde{o}_{t+1} + (1 - \gamma) \tilde{o}_{t-1} - \delta(R_t - E_t \pi_{t+1}) + \eta_{\tilde{y},t} \quad (26)$$

$$\pi_t = \frac{\beta}{1 + \beta\alpha} E_t \pi_{t+1} + \frac{\alpha}{1 + \beta\alpha} \pi_{t-1} + \kappa \tilde{o}_t + \eta_{\pi,t} \quad (27)$$

$$R_t = \rho R_{t-1} + (1 - \rho)(\varphi_\pi \pi_t + \varphi_{\tilde{o}} \tilde{o}_t) + \eta_{R,t} \quad (28)$$

where

$$\eta_{x,t} = \rho_x \eta_{x,t-1} + \omega_{x,t} \quad , \quad -1 < \rho_x < 1 \quad , \quad \omega_{x,t} \sim \text{WN}(0, \sigma_x^2) \quad , \quad x = \tilde{y}, \pi, R. \quad (29)$$

The variables $\tilde{o}_t := o_t - o_t^p$, π_t , and R_t stand for the output gap (o_t is output and o_t^p the natural rate of output), inflation, and the nominal interest rate, respectively; γ is the weight of the forward-looking component in the intertemporal IS curve; α is price the 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 $\beta := 0.99$; κ is the slope of the Phillips curve; ρ , φ_π , and $\varphi_{\tilde{y}}$ 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, $\eta_{\tilde{o},t}$, $\eta_{\pi,t}$ and $\eta_{R,t}$ in Eq. (29) are the mutually independent, autoregressive of order one disturbances and $\omega_{\tilde{o},t}$, $\omega_{\pi,t}$ and $\omega_{R,t}$ are the structural (fundamental) shocks with variances σ_x^2 , $x = \tilde{o}, \pi, R$.

This and 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) and Benati and Surico (2009) have replicated the U.S. Great Moderation, while Castelnuovo and Fanelli (2015) have tested the determinacy/indeterminacy properties of the implied equilibria controlling for identification failure. It is worth noting that Benati and Surico's (2009) model is 'hybrid', in the sense that given the policy rule, both the IS curve and the NKPC feature lags of \tilde{o}_t and π_t other than future expectations. In this respect, it seems particularly suited to serve as a reference structural model in the estimation/evaluation exercise with which we are concerned in this paper.

The three-equation system (26)-(28) can be cast in the form in Eq.s (10)-(11) by setting $Z_t := (\tilde{o}_t, \pi_t, R_t)'$, ($n = 3$), $\eta_t := (\eta_{\tilde{o},t}, \eta_{\pi,t}, \eta_{R,t})'$, $\omega_t := (\omega_{\tilde{o},t}, \omega_{\pi,t}, \omega_{R,t})'$ and

$$\Gamma_0 := \begin{pmatrix} 1 & 0 & \delta \\ -\kappa & 1 & 0 \\ -(1-\rho)\varphi_{\tilde{o}} & -(1-\rho)\varphi_{\pi} & 1 \end{pmatrix}, \Gamma_f := \begin{pmatrix} \gamma & \delta & 0 \\ 0 & \frac{\beta}{1+\beta\alpha} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \Gamma_b := \begin{pmatrix} 1-\gamma & 0 & 0 \\ 0 & \frac{\alpha}{1+\beta\alpha} & 0 \\ 0 & 0 & \rho \end{pmatrix}.$$

$$R := dg(\rho_{\tilde{o}}, \rho_{\pi}, \rho_R) \quad , \quad \Sigma_{\omega} := dg(\sigma_{\tilde{o}}^2, \sigma_{\pi}^2, \sigma_R^2),$$

where the operator $dg(\cdot)$ denotes a diagonal matrix and the entries are in the argument. $\theta := (\gamma, \delta, \alpha, \kappa, \rho, \varphi_{\tilde{y}}, \varphi_{\pi}, \rho_{\tilde{o}}, \rho_{\pi}, \rho_R)$ is the 13×1 vector of structural parameters. The constant C is, in this case, set to zero because estimation is based on demeaned variables, see below.

As in Bårdsen and Fanelli (2015), we complete the model in Eq.s (26)-(29) by assuming that the natural rate of output o_t^p is generated by the Random Walk process:

$$o_t^p = o_{t-1}^p + \eta_{o^p,t} \quad , \quad \eta_{o^p,t} \sim \text{WN}(0, \sigma_{o^p}^2) \quad (30)$$

to capture the effects of technology shocks. Using Eq. (30) and the definition of $\tilde{o}_t := o_t - o_t^p$, we obtain the relationship

$$\tilde{o}_t - \tilde{o}_{t-1} = \Delta o_t - \eta_{o^p,t} \quad (31)$$

where $\Delta o_t := o_t - o_{t-1}$, which will be exploited in the measurement system below.

5.2 Data and statistical model

We employ quarterly data relative to the 'Great Moderation' sample 1984q2-2008q3. The starting date of our estimation and evaluation sample, 1984q2, is justified by McConnell and Pérez-Quirós (2000), who find a break in the variance of the U.S. output growth in 1984q1. The ending date is instead motivated by the fact that, with data after 2008q3, it would be hard to identify

a ‘conventional’ monetary policy shock with our structural model during the well known zero lower bound (ZLB) episodes. We have three observable variables, $y_t := (\Delta o_t, \pi_t, R_t)'$ ($p = 3$), where Δo_t is related to the unobservable output-gap \tilde{o}_t through Eq. (31).¹⁰ Output, o_t , is the log of real GDP. The inflation rate, π_t , is the quarterly growth rate of the GDP deflator. For the short-term nominal interest rate, R_t , we consider the effective Federal funds rate expressed in quarterly terms (averages of monthly values). The source of the data is the website of the Federal Reserve Bank of St. Louis. The three variables are demeaned.¹¹

Step 1 of our estimation and evaluation procedure (Section 4) requires fitting the state-space model in Eq.s (18)-(19) to the data. This entails selecting the optimal length of the state vector, i.e. $\dim(x_t^*) = nk^{op}$. Starting from a maximum lag order of $k^{\max} = 6$, the ‘largest’ statistical model which is taken to the data is given by

$$\begin{pmatrix} Z_t \\ Z_{t-1} \\ \vdots \\ Z_{t-(k^{\max}-1)} \\ x_t^* \end{pmatrix} = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{k^{\max}-1} & \Phi_{k^{\max}} \\ I_3 & 0_{3 \times 3} & \cdots & 0_{n \times n} & 0_{3 \times 3} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0_{3 \times 3} & 0_{3 \times 3} & & I_3 & 0_{3 \times 3} \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-k^{\max}} \\ x_{t-1}^* \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0_{3 \times 3} \\ \vdots \\ 0_{3 \times 3} \\ G^*(\tau) \end{pmatrix} \quad (32)$$

$$\begin{pmatrix} \Delta o_t \\ \pi_t \\ R_t \\ y_t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \tilde{o}_t \\ \pi_t \\ R_t \\ \vdots \\ \tilde{o}_{t-(k^{\max}-1)} \\ \pi_{t-(k^{\max}-1)} \\ R_{t-(k^{\max}-1)} \\ x_t^* \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ v_{1,t} \\ v_t \end{pmatrix} \quad (33)$$

where $Z_t := (\tilde{o}_t, \pi_t, R_t)'$, $\omega_t := (\omega_{\tilde{o},t}, \omega_{\pi,t}, \omega_{R,t})'$, and $v_t := v_{1,t} = \eta_{op,t}$ from Eq. (31).

We estimate the space state model in Eq.s (32)-(33) in the period 1984q2-2008q3, varying k from 1 to $6 =: k^{\max}$, using a Kalman filter-based maximum likelihood approach in conjunction with the simulated annealing/genetic algorithm of Andreassen (2010). For each estimated

¹⁰We have also considered the case in which \tilde{o}_t is proxied by a measure of the output-gap computed by using the measure of the natural rate of output released by the Congressional Budget Office (CBO). In that case, estimation does not necessarily require the use of Eq. (31). Results are available upon request to the authors.

¹¹Before demeaning and estimating the model, we run a preliminary check for stationarity of $y_t := (\Delta o_t, \pi_t, R_t)'$. We compute Johansen’s (1996) cointegration rank test using a VAR model for y_t with restricted (to the cointegration space) and unrestricted constants, respectively. In both cases we reject the hypothesis that unit roots are present in the system. Results are available upon request to the authors.

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 estimate $\hat{\tau}$ delivered by the estimation algorithm. We then compute the Akaike, Hannan-Quinn and Schwartz information criteria and likelihood-ratio (LR) tests to select the optimal lag k^{op} . The results of this specification analysis are summarized in Table 1.

Table 1 shows that using the 5% nominal level of significance, the LR tests selects the model based on $k = k^{op}=4$ lags. The Akaike information criterion selects 5 lags, while Schwartz and Hannan-Quinn select 2 lags. Since the evidence on the optimal number of lags is not clear-cut, it makes sense to consider both of the cases based on $k^{op}=4$ lags, i.e. $\dim(x_t^*)=nk^{op} = 12$, and the case of RE based on $k^{op}=2$ lags, in the period 1984q2-2008q3. In this first case, the vector of parameters associated with the statistical model is given by $\tau := (vec(\Phi_1)', vec(\Phi_2)', vec(\Phi_3)', vec(\Phi_4)', vech(\Sigma_\varepsilon)', \sigma_{\tilde{o}_t}^2)$.

5.3 Frequentist estimation and empirical evaluation

Assuming that the ‘best fitting’ statistical model for the data is based on $k^{op}=4$ lags, i.e. $\dim(x_t^*)=nk^{op} = 12 > \dim(x_t)=n2 = 6$, the pseudo-structural form associated with our NK-DSGE model is given by

$$\tilde{o}_t = \gamma E_t \tilde{o}_{t+1} + (1 - \gamma) \tilde{o}_{t-1} - \delta(R_t - E_t \pi_{t+1}) + \zeta_{\tilde{o},2} \tilde{o}_{t-2} + \zeta_{\tilde{o},3} \tilde{o}_{t-3} + \eta_{\tilde{y},t} \quad (34)$$

$$\pi_t = \frac{\beta}{1 + \beta\alpha} E_t \pi_{t+1} + \frac{\alpha}{1 + \beta\alpha} \pi_{t-1} + \kappa \tilde{o}_t + \zeta_{\pi,2} \pi_{t-2} + \zeta_{\pi,3} \pi_{t-3} + \eta_{\pi,t}$$

$$R_t = \rho R_{t-1} + (1 - \rho)(\varphi_\pi \pi_t + \varphi_{\tilde{o}} \tilde{o}_t) + \zeta_{R,2} R_{t-2} + \zeta_{R,3} R_{t-3} + \eta_{R,t}$$

$$\eta_{x,t} = \rho_x \eta_{x,t-1} + \omega_{x,t} \quad , \quad -1 < \rho_x < 1 \quad , \quad \omega_{x,t} \sim \text{WN}(0, \sigma_x^2) \quad , \quad x = \tilde{o}, \pi, R \quad (35)$$

where $\zeta_{\tilde{o},2}$, $\zeta_{\pi,2}$, $\zeta_{R,2}$, $\zeta_{\tilde{o},3}$, $\zeta_{\pi,3}$ and $\zeta_{R,3}$ are the expectations correction parameters that enter the (supposed diagonal) matrices Υ_2 and Υ_3 , see Eq. (24), in order to rectify the baseline NK-DSGE model. Thus $\zeta := (\zeta_{\tilde{o},2}, \zeta_{\pi,2}, \zeta_{R,2}, \zeta_{\tilde{o},3}, \zeta_{\pi,3}, \zeta_{R,3})' = (diag(\Upsilon_2)', diag(\Upsilon_3)')'$ and $\theta^* = (\theta', \zeta')' := (\gamma, \delta, \alpha, \kappa, \rho, \varphi_{\tilde{o}}, \varphi_\pi, \rho_{\tilde{o}}, \rho_\pi, \rho_R, \sigma_{\tilde{o}}^2, \sigma_\pi^2, \sigma_R^2, \zeta_{\tilde{o},2}, \zeta_{\pi,2}, \zeta_{R,2}, \zeta_{\tilde{o},3}, \zeta_{\pi,3}, \zeta_{R,3})'$ is the 19×1 vector containing the truly structural and expectations correction parameters.

Step 2 of the procedure summarized in Section 4 requires estimating $\theta^* = (\theta', \zeta')'$ from the state space model (32)-(33) by imposing the CER implied by the pseudo-structural form. By applying Proposition 1, we obtain: $\Phi_i = \tilde{\Phi}_i$, $i = 1, 2, 3, 4$, $\Sigma_\varepsilon = \tilde{\Sigma}_\varepsilon$, where

$$\left\{ \begin{array}{l} (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_1 - (\Gamma_f \tilde{\Phi}_2 + \Gamma_{b,1}) = 0_{3 \times 3} \\ (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_2 - (\Gamma_f \tilde{\Phi}_3 + \Upsilon_2 - R \Gamma_b) = 0_{3 \times 3} \\ (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_3 - (\Gamma_f \tilde{\Phi}_4 + \Upsilon_3 - R \Upsilon_2) = 0_{3 \times 3} \\ (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1) \tilde{\Phi}_4 + R \Upsilon_3 = 0_{3 \times 3} \\ \tilde{\Sigma}_\varepsilon - \tilde{\Psi} \Sigma_\omega \tilde{\Psi}' = 0_{n \times n} \quad , \quad \tilde{\Psi} := (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1)^{-1}. \end{array} \right. \quad (36)$$

Estimation results for $\theta^* = (\theta', \zeta')'$ are reported in Table 2. In the upper panel of Table 2, we summarize the estimate of θ obtained under RE, i.e. taking the structural model in Eq.s (26)-(29) to the data in the ‘conventional way’, and the corresponding estimate obtained from the pseudo-structural form (34)-(35). Observe that in order to obtain the CER under RE ($k^{op}=2$), it is sufficient to set the expectations correction matrices Υ_2 and Υ_3 to zero in Eq. (36), leading to $\tilde{\Phi}_3 = \tilde{\Phi}_4 = 0_{3 \times 3}$. This implies that there are no deviations of agents’ expectations from RE. We label the estimates obtained from the pseudo-structural form in the fourth column of Table 2 with the acronym ‘QRE’.

The lower panel of Table 2 summarizes the Akaike, Hannan-Quinn and Schwartz information criteria and a battery of LR tests through which it is possible to select the ‘best’ specification. All three information criteria favour the model estimated under QRE. The LR test for the null $\zeta=0_{6 \times 1}$ (RE) against the alternative $\zeta \neq 0_{6 \times 1}$ (QRE) strongly rejects the null hypothesis. In both cases, the CER are strongly rejected.

Coming back to the estimated parameters in the upper panel of Table 2, we notice that the large majority (four out of six) of the expectations correction parameters $\zeta := (\zeta_{\delta,2}, \zeta_{\pi,2}, \zeta_{R,2}, \zeta_{\delta,3}, \zeta_{\pi,3}, \zeta_{R,3})'$, reported in the forth column are significant at conventional significance levels. This confirms that there is a mismatch between agents’ expectations and the case of RE. Focusing on the truly structural parameters θ , we notice that the main differences between the estimates obtained under RE and QRE involve the intertemporal elasticity of substitution δ and the forward-looking parameter γ in the IS curve, the slope κ and shock persistence parameter ρ_π in the NKPC, and the Fed’s long run response to output gap φ_δ and shock persistence parameter ρ_π in the policy reaction function.

The magnitude and precision of the estimated δ is considerably higher under QRE, whereas δ does not seem to be empirically identified under RE. Conversely, the magnitude and precision of the estimated γ is lower under QRE relative to RE, suggesting a lesser extent of forward-looking behaviour once we account for the whole dynamics of the system. This result can be clearly explained in light of the more ‘flexible’ expectations generating system we assume. The slope parameter of the NKPC is poorly estimated in both cases, confirming a traditional difficulty in its empirical identification. The magnitude of the estimated indexation parameter of the NKPC, α , is the same in the two cases; we observe that precision is considerably higher under QRE relative to RE. However, the estimated α obtained under QRE is comparatively more precise than the estimate obtained under RE. Overall, our maximum likelihood estimates seems to suggest that the NKPC can be more precisely empirically identified by relaxing some constraints on the autocorrelation structure of the data. As concerns the policy rule, we notice that the Fed’s long run response to output gap is remarkably higher relative to the case of RE

(1.5 as opposed to 0.336) and more precisely estimated under QRE. As it known, the empirical literature on the identifiability of the policy parameters $\varphi_{\tilde{o}}$ and φ_{π} in the class of New Keynesian models is huge and has not yet reached a consensus. The recent empirical literature, which makes increasing use of identification-robust methods, suggests that it is difficult to estimate $\varphi_{\tilde{o}}$ (and φ_{π}) precisely on the Great Moderation era, see, among others, Mavroeidis (2010) and Castelnuovo and Fanelli (2015) and references therein. Our result seems to suggest that concern regarding the dynamic misspecification that characterize baseline New Keynesian systems can aid the empirical identification process of monetary policy parameters.

5.4 Bayesian estimation and empirical evaluation

In the Bayesian approach, Step 1 is exactly as in Sub-section 5.2; hence, the estimated pseudo-structural form in Step 2 is given by system (34)-(35). The priors used for the truly structural parameters, θ , are taken from Benati and Surico (2009), while the priors used for the expectations correction parameters, ζ , are centered at the RE equilibrium. This means that for each $\zeta_{i,j}$, $i = \tilde{o}, \pi, R$, $j = 2, 3$ in Eq.s (34)-(35), we use a Gaussian distribution centered on 0 with variance 0.25. Table 3 summarizes the modes and standard deviations of the prior distributions for all structural parameters. The RWM algorithm delivers the posterior distributions reported in Table 4.

As expected, the DIC information criterion favours the NK-DSGE model estimated under QRE, relative to the case of RE. The estimates in Table 4 are quantitatively different from their counterparts in Table 2 obtained with the frequentist maximum likelihood approach. Similarly to the frequentist estimation approach, we observe that the mismatch between agents' expectations and RE seems to be relevant. The magnitude of estimated persistence parameters, $\rho_{\tilde{o}}$, ρ_{π} and ρ_R , is considerably larger in the pseudo-structural form compared to the case of RE, suggesting that other than capturing omitted propagation mechanisms, the pseudo-structural model does not penalize the persistence of the data, given the chosen priors.

The main differences between the estimates obtained under RE and by the pseudo-structural form involve the forward-looking parameter of the IS curve, γ , and the policy reaction of the Fed to the output gap, $\varphi_{\tilde{o}}$. Contrary to what is reported in Table 2, the magnitude of the estimated γ is considerably larger under QRE, pointing towards a greater extent of forward-looking behaviour. Obviously, the difference in the estimates of γ in Table 2 and Table 4 can solely be ascribed to the role of the prior distributions. On the other hand, we notice that in the Bayesian approach as well, the Fed's long run response to output gap is remarkably higher relative to the case of RE (1.054 as opposed to 0.449) and more precisely estimated under QRE. This evidence confirms the finding obtained with the frequentist approach in Sub-section 5.3.

All other estimates in Table 4 are roughly the same as in Table 2.

6 Concluding remarks

In this paper, we have focused on the poor time series performance that characterizes the class of small-scale NK-DSGE models used in monetary policy and business cycle analysis. Under RE, NK-DSGE models give rise to a set of nonlinear CER and constraints on the lag order of the system that may conflict with the actual autocorrelation structure that characterizes quarterly (or monthly) time series. In these cases, the investigator should re-formulate the structural model by specifying a less restrictive, possibly microfounded, dynamic structure that accounts for the previously omitted propagation mechanisms. This is not always feasible. Practitioners typically react to the misspecification issue by specifying *ad hoc* time-series models for the shocks in order to improve the overall fit of the model. We rationalize these practices using a statistical model for the data that is treated like the true agents' expectations generating mechanism. This allows one to relax the tightness of the restrictions arising under RE, without abandoning the logic of model-consistent expectations. Our approach is illustrated empirically focusing on the 'hybrid' NK-DSGE monetary model by Benati and Surico (2009) as the reference system.

A Appendix: proof of propositions

Proof of Proposition 1 For $k^{op} \leq 2$, the result is obtained under Assumptions 1-5 by following Binder and Pesaran (1995) and, more in detail, the supplementary material in Castelnuovo and Fanelli (2015). For $k^{op} \geq 3$, use the autoregressive structure of η_t to write system (24) in the form

$$\Gamma_0^R Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_{b,1} Z_{t-1} + \Upsilon_2^* Z_{t-2} + \sum_{j=3}^{k^{op}-1} \Upsilon_j^* Z_{t-j} + \Upsilon_{k^{op}}^* Z_{t-k^{op}} + C^* + \omega_t^*$$

where $\omega_t^* := \omega_t + R\Gamma_f u_t$ and $u_t := Z_t - E_{t-1} Z_t$ is a martingale difference sequence, $\Gamma_0^R := (\Gamma_0 + R\Gamma_f)$, $\Gamma_{b,1} := (\Gamma_b + R\Gamma_0)$, $\Upsilon_2^* := (\Upsilon_2 - R\Gamma_b)$, $\Upsilon_j^* := (\Upsilon_j - R\Upsilon_{j-1})$, $j = 3, \dots, k^{op} - 1$, $\Upsilon_{k^{op}}^* := -R\Upsilon_{k^{op}-1}$ and $C^* := (I_n - R)C$. Then compact the system above in the companion-form representation:

$$\Gamma_0^c Z_t^c = \Gamma_f^c E_t Z_{t+1}^c + \Gamma_b^c Z_{t-1}^c + C^c + \omega_t^c \quad (37)$$

where $Z_t^c := (Z_t', Z_{t-1}', \dots, Z_{t-k^{op}+1}')'$, $C^c := (C^{*'}, 0_{1 \times n(k^{op}-1)})'$, $\omega_t^c := (\omega_t^{*'}, 0_{1 \times n(k^{op}-1)})'$ and

$$\Gamma_0^c := \begin{pmatrix} \Gamma_0^R & 0_{n \times n} & \cdots & 0_{n \times n} \\ 0_{n \times n} & I_n & \cdots & 0_{n \times n} \\ \vdots & \vdots & \ddots & \\ 0_{n \times n} & 0_{n \times n} & \cdots & I_n \end{pmatrix}, \quad \Gamma_f^c := \begin{pmatrix} \Gamma_f & 0_{n \times n} & \cdots & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} & \cdots & 0_{n \times n} \\ \vdots & \vdots & \ddots & \\ 0_{n \times n} & 0_{n \times n} & \cdots & 0_{n \times n} \end{pmatrix}$$

$$\Gamma_b^c := \begin{pmatrix} \Gamma_{b,1} & \Upsilon_2^* & \Upsilon_3^* & \cdots & \Upsilon_{k^{op}-1}^* \\ I_n & 0_{n \times n} & 0_{n \times n} & \cdots & 0_{n \times n} \\ \vdots & \ddots & & \vdots & \vdots \\ 0_{n \times n} & 0_{n \times n} & 0_{n \times n} & I_n & 0_{n \times n} \end{pmatrix}.$$

From Binder and Pesaran (1995) and the supplementary material in Castelnovo and Fanelli (2014), it follows that if under Assumptions 1-4 a unique and stable solution for the system (37) exists, it takes the form

$$Z_t^c = \tilde{\Phi}^c Z_{t-1}^c + \tilde{\mu}^c + \varepsilon_t^c, \quad \varepsilon_t^c := \tilde{\Psi}^c \omega_t^c$$

where $\tilde{\Phi}^c$, $\tilde{\Psi}^c$ and $\tilde{\mu}^c := (\mu', 0_{1 \times n(k^{op}-1)})'$ are block matrices (vectors) subject to the CER:

$$(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c) \tilde{\Phi}^c = \Gamma_b^c = 0_{nk^{op} \times nk^{op}} \quad (38)$$

$$(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c - \Gamma_f^c) \tilde{\mu}^c = C^c \quad (39)$$

$$E(\varepsilon_t^c \varepsilon_t^{c'}) := \tilde{\Psi}^c E(\omega_t^c \omega_t^{c'}) \tilde{\Psi}^{c'}, \quad \tilde{\Psi}^c := (\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c)^{-1}$$

and with $k = k^{op}$ the matrix $\tilde{\Phi}^c$ has the same block structure as the matrix $A(\tau)$ in Eq. (18). Assumption 1 ensures that $\tilde{\Phi}^c$ is stable, while uniqueness occurs if $(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c)^{-1} \Gamma_f^c$ is stable. Note that

$$(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c) := \begin{pmatrix} \Gamma_0^R - \Gamma_f \tilde{\Phi}_1 & -\Gamma_f \tilde{\Phi}_2 & \cdots & -\Gamma_f \tilde{\Phi}_{k^{op}} \\ 0_{n \times n} & I_n & \cdots & 0_{n \times n} \\ \vdots & \vdots & \ddots & \\ 0_{n \times n} & 0_{n \times n} & \cdots & I_n \end{pmatrix}$$

and using inversion formulas for partitioned matrices, the upper left block of the inverse of the matrix above ($\tilde{\Psi}^c$) is $\tilde{\Psi} := (\Gamma_0^R - \Gamma_f \tilde{\Phi})^{-1}$. Moreover,

$$(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c)^{-1} \Gamma_f^c := \begin{pmatrix} (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1)^{-1} \Gamma_f & 0_{n \times n} & \cdots & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} & \cdots & 0_{n \times n} \\ \vdots & \vdots & \ddots & \\ 0_{n \times n} & 0_{n \times n} & \cdots & 0_{n \times n} \end{pmatrix}$$

and

$$(\Gamma_0^c - \Gamma_f^c \tilde{\Phi}^c - \Gamma_f^c) \tilde{\mu}^c - C^c := \begin{pmatrix} (\Gamma_0^R - \Gamma_f \tilde{\Phi}_1 - \Gamma_f) \tilde{\mu} - (I_n - R)C \\ 0_{n \times n} \\ \vdots \\ 0_{n \times n} \end{pmatrix}$$

Using simple algebra it turns out that the relationships in Eqs. (38)-(39) are equivalent to the CER in Eq. (25).■

References

- An, S. and 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, 127-154.
- Andrews, I. and Mikusheva, A. (2014): Maximum likelihood inference in weakly identified DSGE models, *Quantitative Economics*, forthcoming.
- Bårdsen, G. and Fanelli, L. (2015): Frequentist evaluation of small DSGE models. *Journal of Business and Economic Statistics*, forthcoming.
- Benati L. and Surico, P. (2009): VAR analysis and the Great Moderation, *American Economic Review* 99, 1636-1652.
- Beyer, A., Farmer, R., Henry, J. and Marcellino, M. (2008): Factor analysis in a model with rational expectations, *The Econometrics Journal* 11, 271-286.
- Binder, M. and Pesaran, M. H. (1995): Multivariate rational expectations models and macroeconomic modelling: a review and some new results. In M. H. Pesaran and M. Wickens (eds.), *Handbook of Applied Econometrics*, pp. 139-187 (Ch. 3). Oxford: Blackwell.
- Boivin, J. and Giannoni, M.P. (2006): Has monetary policy become more effective ?, *Review of Economics and Statistics* 88, 445-462.
- Branch, W.A. and Evans, G.W. (2006): A simple recursive forecasting model, *Economic Letters* 91, 158-166.

- Canova, F. and Ferroni, F. (2012): The dynamics of US inflation: Can monetary policy explain the changes? *Journal of Econometrics* 167, 47-60.
- Canova, F. and Sala, L. (2009): Back to square one: identification issues in DSGE models. *Journal of Monetary Economics* 56, 431-449.
- Castelnuovo, E. and Fanelli, L. (2015): Monetary policy indeterminacy and identification failures in the U.S.: Results from a robust test, *Journal of Applied Econometrics*, forthcoming.
- Chevillon, G., Massmann, M. and Mavroeidis, S. (2010): Inference in models with adaptive learning, *Journal of Monetary Economics* 57, 341-351.
- Cho I-K and Kasa, K. (2014): Learning and model validation, *Review of Economic Studies*, forthcoming.
- Clarida R.J., Gali, J. and Gertler, M. (2000): Monetary policy rules and macroeconomic stability: evidence and some theory, *Quarterly Journal of Economics* 115, 147-180.
- Cole, S. J. and Milani, F. (2014): The Misspecification of Expectations in New Keynesian Models: A DSGE-VAR Approach. Working paper.
- Consolo, A., Favero, C.A. and Paccagnini, A. (2009): On the statistical identification of DSGE models, *Journal of Econometrics* 150, 99-115.
- Cúrdia, V. and Reis, R. (2010): Correlated disturbances and U.S. business cycle, *NBER Working Paper* 15774.
- Del Negro, M., Schorfheide, F., Smets, F. and Wouters, R. (2007): On the fit of New Keynesian models, *Journal of Business and Economic Statistics* 25, 123-143.
- Diebold, F. X., Ohanian, L. E. and Berkowitz, J. (1998): Dynamic equilibrium economies: a framework for comparing models and data, *Review of Economic Studies* 65, 433-452.
- Dufour J-M, Khalaf L and Kichian M. (2013): Identification-robust analysis of DSGE and structural macroeconomic models. *Journal of Monetary Economics* 60, 340-350.
- Evans, G. W. and Honkapohja, S. (1999): *Learning dynamics*, in Handbook of Macroeconomics 1A, Chap. 7
- Evans, G. W. and Honkapohja, S. (2001): *Learning and expectations in macroeconomics*, Princeton University Press.

- Evans, G. W. and Honkapohja, S. (2003): Adaptive learning and monetary policy design. *Journal of Money, Credit and Banking*, 1045-1072.
- Fanelli, L. (2008): Evaluating the New Keynesian Phillips Curve under VAR-based Learning, *Economics: The Open-Access, Open-Assessment E-Journal* 2, 2008-33.
- Fanelli, L. (2009): Estimation of Quasi-Rational DSGE monetary models, Dipartimento di Scienze Statistiche, Quaderno di Dipartimento Serie Ricerche 2009 n.3.
- Fanelli, L. (2012): Determinacy, indeterminacy and dynamic misspecification in Linear Rational Expectations models, *Journal of Econometrics* 170, 153-163.
- Fanelli, L. and Palomba, G. (2011): Simulation-based tests of forward-looking models under VAR learning dynamics, *Journal of Applied Econometrics* 26, 762-782.
- Fernández-Villaverde, J., Rubio-Ramírez, J., Sargent, T. and Watson, M. (2007): ABCs (and Ds) of understanding VARs, *American Economic Review* 97, 1021-1026.
- Franchi, M. and Vidotto, A. (2013): A check for finite order VAR representations of DSGE models, *Economic Letters* 120, 100-103.
- Franchi, M. and Paruolo, P. (2014): Minimality of state space solutions of DSGE models and existence conditions for their VAR representation, *Computational Economics*, forthcoming.
- Franchi, M. and Juselius, K. (2007): Taking a DSGE model to the data meaningfully, *Economics: The Open-Access, Open-Assessment E-Journal* 1, 2007-4
- Fuhrer, J. and Rudebusch, G.D. (2004): Estimating the Euler equation for output, *Journal of Monetary Economics* 51, 1133-1153.
- Fukač, M. and Pagan, A. (2009): Limited information estimation and evaluation of DSGE models, *Journal of Applied Econometrics*, 25, 55-70.
- Guerron-Quintana P., Inoue, A. and Kilian, L. (2013): Frequentist inference in weakly identified DSGE models, *Quantitative Economics* 4, 197-229.
- Hannan, E.J. and Deistler, M. (1988): *The Statistical Theory of Linear Systems*, SIAM ed. Wiley, New York.
- Hansen, L.P. (2014): Nobel Lecture: Uncertainty outside and inside economic models, *Journal of Political Economy*, 122, 945-987.

- Hansen, L. P. and Sargent, T. J. (1980): Formulating and estimating dynamic linear rational expectations models, *Journal of Economic Dynamics and Control* 2, 7-46.
- Hansen, L. P. and Sargent, T. J. (1981): Linear rational expectations models for dynamically interrelated variables, in: R.E.Jr. Lucas and T. J. Sargent, (Eds.), *Rational Expectations and Econometric Practise*. University of Minnesota Press, Minneapolis, pp. 127-156.
- Henry, S.G.B. and Pagan, A.R. (2004): The econometrics of the New Keynesian policy model: Introduction, *Oxford Bulletin of Economics and Statistics* 66, Supplement, 581-607.
- Holt, M.T. and 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.
- Johansen, S. (1996): *Likelihood-based inference in cointegrated vector auto-regressive models*, Oxford University Press, Oxford.
- Jondeau, E. and Le Bihan, H. (2008): Examining bias in estimators of linear rational expectations models under misspecification, *Journal of Econometrics* 143, 375-395.
- Klein, P. (2000): Using the generalized Schur form to solve a multivariate linear rational expectations model, *Journal of Economic Dynamics and Control* 24, 1405-1423.
- Komunjer, I. and Ng, S. (2011): Dynamic identification of dynamic stochastic general equilibrium models, *Econometrica* 79, 1995-2032.
- Kozicki, S. and Tinsley, P. (1999): Vector rational error correction, *Journal of Economic Dynamics and Control* 23, 1299-1327.
- Lubik, T.A. and Schorfheide, F. (2004): Testing for indeterminacy: an application to U.S. monetary policy, *American Economic Review* 94, 190-217.
- Mavroeidis, S. (2010): Monetary policy rules and macroeconomic stability: some new evidence, *American Economic Review* 100, 491-503.
- McConnell M.M. and Pérez-Quirós, G. (2000): Output fluctuations in the United States: what has changed since the early 1980's? *American Economic Review* 90, 1464-1476.
- Milani, F. (2007): Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics* 54, 2065-2082.
- Nelson, R. G. and Bessler, D. A. (1992): Quasi-rational expectations: Experimental evidence, *Journal of Forecasting* 11, 141-156.

- Nerlove, M. and 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. and Carvalho, J.L. (1979): *Analysis of Economic Time Series*. Academic Press, New York.
- Guerron-Quintana, P., Inoue, A. and Kilian, L. (2013): Frequentist inference in weakly identified DSGE models, *Quantitative Economics* 4, 197-229.
- 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 monetary 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.
- Sargent, T.J. (1999): The conquest of American inflation. *Princeton University Press*, Princeton.
- Smets, F. and Wouters, R. (2007): Shocks and frictions in U.S. business cycles, *American economic Review* 97, 586-606.
- Uhlig, H. (1999): A Toolkit for analyzing nonlinear dynamic stochastic models easily. In R. Marimon and A. Scott (eds.), *Computational methods for the study of dynamics economies*, pp. 30-61, Oxford University Press, New York.

TABLE 1. Lag length selection in the statistical model in Eq.s (32)-(33).

Estimation sample: 1984q2 - 2008q3						
	LR tests			Information criteria		
lag	Likelihood	LR	p-value	Akaike	Hannan-Quinn	Schwartz
2	151.42	74.71	0.000	-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 Andreassen (2010). The LR tests are computed by comparing the log-likelihoods obtained with $k = 2, \dots, 5 =: k^{\max} - 1$ lags with the log-likelihood obtained with $k^{\max} = 6$. Asterisks denote the optimal lag selection according to the test/information criterion.

TABLE 2. Estimated structural parameters of the Benati and Surico's (2009) model in Eq.s (26)-(29).

Estimation sample: 1984q2 - 2008q3			
Parameters	Interpretation	RE	QRE
δ	IS: inter. elast. of substitution	0.010(0.057)	0.079(0.055)
γ	IS: forward looking term	0.572(0.062)	0.269(0.207)
α	NKPC: indexation past inflation	0.035(0.230)	0.035(0.039)
κ	NKPC: slope	0.041(0.121)	0.0267(0.043)
ρ	Policy rule: smoothing term	0.908(0.054)	0.889(0.034)
$\varphi_{\bar{o}}$	Policy rule: reaction to output gap	0.336(0.963)	1.500(0.248)
φ_{π}	Policy rule: reaction to inflation	1.650(0.974)	1.650(0.803)
$\rho_{\bar{o}}$	IS: shock persistence	0.908(0.034)	0.801(0.190)
ρ_{π}	NKPC: shock persistence	0.100(0.342)	0.775(0.082)
ρ_R	Policy rule: shock persistence	0.539(0.080)	0.192(0.157)
$\sigma_{\bar{o}}^2$	IS: variance of shock	0.001(0.001)	0.006(0.002)
σ_{π}^2	NKPC: variance of shock	0.025(0.003)	0.053(0.010)
σ_R^2	Policy rule: variance of shock	0.011(0.002)	0.006(0.001)
σ_{op}^2	Variance of potential output	0.045(0.009)	0.031(0.006)
$\zeta_{\bar{o},2}$	IS: exp. correction (Υ_2)	-	-0.061(0.191)
$\zeta_{\pi,2}$	NKPC: exp. correction (Υ_2)	-	-0.444(0.176)
$\zeta_{R,2}$	Policy rule: exp. correction (Υ_2)	-	0.057(0.061)
$\zeta_{\bar{o},3}$	IS: exp. correction (Υ_3)	-	0.047(0.016)
$\zeta_{\pi,3}$	NKPC: exp. correction (Υ_3)	-	0.065(0.131)
$\zeta_{R,3}$	Policy rule: exp. correction (Υ_3)	-	-0.192(0.058)
Likelihood		115.09	129.79
Akaike		-202.18	-219.57*
Hannan-Quinn		-187.54	-198.66*
Schwartz		-165.99	-167.88*

LR(RE vs QRE)=29.40, p-value=0.000

LR(CER model with RE)=72.66, p-value=0.000

LR(CER model with QRE)=94.40, p-value=0.0001

NOTES: 'RE' means that the model in Eq.s (26)-(29) is estimated in the 'conventional' way under rational expectations; 'QRE' means that the estimated model is the pseudo-structural form in Eq.s (34)-(35). 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 γ ; [0.035-0.100] for α ; [0.025- ∞] for κ ; [0.001-0.999] for ρ ; [0.001-1.500] for φ_{δ} ; [1.650-5.500] for φ_{π} ; [0.001-0.999] for ρ_{δ} , ρ_{π} and ρ_R , leaving all remaining parameters, including the auxiliary parameters $\zeta := (\zeta_{\delta,2}, \zeta_{\pi,2}, \zeta_{R,2}, \zeta_{\delta,3}, \zeta_{\pi,3}, \zeta_{R,3})'$, free on condition that model's determinacy was met. Standard errors in parentheses have been calculated using the 'hessian.mat' function available in Matlab. Asterisks denote the optimal lag selection according to the information criterion.

TABLE 3. Bayesian approach, prior distributions used for the structural parameters in the Benati and Surico's (2009) model in Eq.s (26)-(29).

Parameter	Interpretation	Density	Mode	Standard Deviation
δ	IS: inter. elast. of substitution	<i>Inverse Gamma</i>	0.06	0.04
γ	IS: forward looking term	<i>Beta</i>	0.25	0.20
α	NKPC: indexation past inflation	<i>Beta</i>	0.75	0.20
κ	NKPC: slope	<i>Gamma</i>	0.05	0.01
ρ	Policy rule: smoothing term	<i>Beta</i>	0.75	0.20
$\varphi_{\tilde{o}}$	Policy rule: reaction to output gap	<i>Gamma</i>	0.15	0.25
φ_{π}	Policy rule: reaction to inflation	<i>Gamma</i>	1.00	0.50
$\rho_{\tilde{o}}$	IS: shock persistence	<i>Beta</i>	0.25	0.20
ρ_{π}	NKPC: shock persistence	<i>Beta</i>	0.25	0.20
ρ_R	Policy rule: shock persistence	<i>Beta</i>	0.25	0.20
$\sigma_{\tilde{o}}^2$	IS: variance of shock	<i>Inverse Gamma</i>	0.25	0.25
σ_{π}^2	NKPC: variance of shock	<i>Inverse Gamma</i>	0.50	0.50
σ_R^2	Policy rule: variance of shock	<i>Inverse Gamma</i>	0.25	0.25
σ_{τ}^2	Variance of potential output	<i>Inverse Gamma</i>	0.25	0.25
$\zeta_{i,j}$	Auxiliary, $i = \tilde{o}, \pi, R$; $j = 2, 3$	<i>Normal</i>	0	0.25

NOTES: The prior distributions for the truly structural parameters, θ , are taken from Table 1 in Benati and Surico (2009). The parameter δ corresponds to σ^{-1} in Benati and Surico (2009), hence we use an Inverse-Gamma distribution in place of a Gamma.

TABLE 4. Bayesian approach, estimated structural parameters of the Benati and Surico's (2009) model in Eq.s (26)-(29).

Estimation sample: 1984q2 - 2008q3			
Parameters	Interpretation	Posterior RE	Posterior QRE
		Mean [5%, 95%]	Mean [5%, 95%]
δ	IS: inter. elast. of substitution	0.183[0.156,0.199]	0.185[0.158,0.199]
γ	IS: forward looking term	0.136[0.102,0.200]	0.829[0.649,0.951]
α	NKPC: indexation past inflation	0.056[0.036,0.088]	0.062[0.037,0.093]
κ	NKPC: slope	0.053[0.038,0.072]	0.053[0.038,0.071]
ρ	Policy rule: smoothing term	0.783[0.682,0.870]	0.733[0.584,0.902]
φ_{δ}	Policy rule: reaction to output gap	0.449[0.116,0.805]	1.054[0.167,1.478]
φ_{π}	Policy rule: reaction to inflation	2.107[1.682,3.043]	1.801[1.658,2.120]
ρ_{δ}	IS: shock persistence	0.529[0.341,0.714]	0.845[0.681,0.952]
ρ_{π}	NKPC: shock persistence	0.484[0.175,0.771]	0.783[0.637,0.903]
ρ_R	Policy rule: shock persistence	0.470[0.204,0.717]	0.641[0.233,0.916]
σ_{δ}^2	IS: variance of shock	0.041[0.030,0.054]	0.047[0.034,0.063]
σ_{π}^2	NKPC: variance of shock	0.232[0.193,0.278]	0.249[0.204,0.301]
σ_R^2	Policy rule: variance of shock	0.112[0.092,0.135]	0.115[0.094,0.139]
$\sigma_{\delta p}^2$	Variance of potential output	0.048[0.035,0.063]	0.054[0.039,0.073]
$\zeta_{\delta,2}$	IS: exp. correction (Υ_2)	-	-0.082[-0.364,0.209]
$\zeta_{\pi,2}$	NKPC: exp. correction (Υ_2)	-	-0.638[-0.913,-0.370]
$\zeta_{R,2}$	Policy rule: exp. correction (Υ_2)	-	0.125[-0.176,0.439]
$\zeta_{\delta,3}$	IS: exp. correction (Υ_3)	-	0.011[-0.249,0.263]
$\zeta_{\pi,3}$	NKPC: exp. correction (Υ_3)	-	-0.247[-0.538,0.055]
$\zeta_{R,3}$	Policy rule: exp. correction (Υ_3)	-	-0.053[-0.307,0.213]
DIC		42.41	3.26*

NOTES: 'RE' means that the model in Eq.s (26)-(29) is estimated in the 'conventional' way under rational expectations; 'QRE' means that the estimated model is the pseudo-structural form in Eq.s (34)-(35). The prior distributions are reported in Table 3. Posterior distributions are computed using the Random Walk Metropolis algorithm. 'Mean [5%, 95%]' denotes a 90% credible set. The posteriors satisfy the standard convergence criteria and the acceptance ratio is 22.94% for the model estimated under RE and 36.79 for the model estimated under QRE. DIC is the Deviance Information Criterion.