Frequentist evaluation of small DSGE models

Gunnar Bårdsen*   Luca Fanelli†

This version: Janaury 2014

Abstract

This paper proposes a new evaluation approach for the class of small-scale 'hybrid' New Keynesian Dynamic Stochastic General Equilibrium (NK-DSGE) models typically used in monetary policy and business cycle analysis. The empirical assessment of the NK-DSGE model is based on a conditional sequence of likelihood-based tests conducted in a Vector Autoregressive (VAR) system, in which both the low and high frequency implications of the model are addressed in a coherent framework. If some of the low frequency behavior of the original time series of the model can be approximated by non-stationary processes, stationarity must be imposed by removing the stochastic trends. This gives rise to a set of recoverable unit roots/cointegration restrictions, in addition to the short-run cross-equation restrictions. The procedure is based on the sequence 'LR1→LR2→LR3', where LR1 is the cointegration rank test, LR2 the cointegration matrix test and LR3 the cross-equation restrictions test: LR2 is computed conditional on LR1 and LR3 is computed conditional on LR2. The type-I errors of the three tests are set consistently with a pre-fixed overall nominal significance level. A bootstrap analogue of the testing strategy is proposed in small samples. We show that the information stemming from the individual tests can be used constructively to uncover which features of the data are not captured by the theoretical model and thus to rectify, when possible, the specification. We investigate the empirical size properties of the proposed testing strategy by a Monte Carlo experiment and show the empirical usefulness of our approach by estimating and testing a monetary business cycle NK-DSGE model using U.S. quarterly data.

Keywords: DSGE models, LR test, Maximum Likelihood, New-Keynesian model, VAR

J.E.L. C5, E4, E5

*Department of Economics, Norwegian University of Science and Technology, Dragvoll, 7491, Trondheim.

†Department of Statistical Sciences and School of Economics, Statistics and Management, University of Bologna, via Belle Arti 41, I-40126 Bologna.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are dominating macroeconomics, in academic research, as well as in economic policy making. Even though these models, by their very nature, cannot provide a complete description of the business cycle and of any time series, such as inflation, output and the policy rate, they are widely used to evaluate macroeconomic scenarios and predict economic activity. Assessing the correspondence between what these models imply and what the data tell us is therefore a crucial step in the process of analyzing policy options and their effects, especially if one takes the view that the scientific validity of a model should not be exclusively based on its logical coherence or its intellectual appeal, but also on its ability to make empirical predictions that are not rejected by the data; see e.g. De Grauwe (2010) and Pesaran and Smith (2011).

There are several methods which can used to evaluate the empirical performance of DSGE models, depending on the specific objectives of the analysis. Most common methods include economic reliability, statistical fit, and forecasting accuracy; see, e.g., Schorfheide (2000), An and Schorfheide (2007) and Schorfheide (2011). Each evaluation method is based on a 'metric' and different 'metrics' may lead to different conclusions. Our 'metric' will be based on testing the restrictions on the data implied by DSGE models. This approach is by no means new, but dates back to the early literature on the econometrics of rational expectations models; see Hansen and Sargent (1980), Hansen and Sargent (1981), Wallis (1980) and Johansen and Swensen (1999).

It is often claimed that Bayesian techniques are preferable to standard likelihood-based methods because DSGE models typically represent a false description of the Data Generating Process (DGP) and misspecification can be important in estimation; see e.g. Canova and Ferroni (2012). Schorfheide (2000) suggests using a loss function to assess the discrepancy between DSGE model predictions and overall posterior distribution of the population characteristics that the researcher is trying to match. Del Negro et al. (2007) develop a set of tools within the Bayesian approach that can be used for assessing the time series fit of a DSGE model based on a systematic relaxation of the set of cross-equation restrictions (CER) that the structural model implies on the Vector Autoregressive (VAR) representation of the data. Their method, known as the 'DSGE-VAR' approach, provides the investigator with a Bayesian 'metric' through which he/she can evaluate how far/close the DSGE model is from a VAR approximation of the data. While misspecification in DSGE models is a concrete possibility, we do not think it represents a strong argument against the idea of confronting these models with data by frequentist (classical) methods. The knowledge that the DSGE model is 'misspecified' in some directions may help the investigator understand what features of the data the model is missing, how important these features are, and, possibly, how to improve the original specification.

We propose a frequentist evaluation approach for a class of small-scale DSGE models grounded in the New Keynesian tradition and relevant for economic policy analysis, henceforth denoted with the acronym 'NK-DSGE' models. These models are investigated in, among many others, Clarida et al. (2000), Lubik and Schorfheide (2004), Ireland (2004), Christiano et al. (2005), Smets and Wouters (2007), DeJong and Dave (2011), Carlstrom et al. (2009), Benati and Surico (2009) and more generally, in Woodford (2003) and Galí (2008) They feature both macroeconomic and monetary policy shocks and typically include a forward-looking aggregate demand equation, a Phillips curve, and a monetary policy reaction function. They can also accommodate the monetary/fiscal policy mix (e.g. Bianchi (2012)) and/or financial frictions e.g. Castelnuovo and Nisticò (2010).

In principle, there are two types of restrictions that can be tested in NK-DSGE models. First, there are the long-run cointegration/common-trend restrictions stemming from the observation that there are generally more variables to be modelled than there are independent integrated forcing processes; see e.g. Canova et al. (1994), Söderlind and Vredin (1996), Fukač and Pagan (2010) and Juselius (2011). Importantly, these restrictions hold regardless of the uniqueness/multiplicity of the model solution; see Broze et al. (1990) and Binder and Pesaran (1995). Hence the restrictions are invariant to the specification of the transient dynamics of the system. Second, there are the short-run CER which apply to the system conditional on the common trends. The long-run and short-run properties of NK-DSGE models are generally interdependent and therefore they should be examined jointly. Our method is based on testing both types of restrictions in a coherent framework. We propose a sequential procedure computed in three steps using likelihood ratio (LR) tests. We first test whether the cointegration rank (the number of stochastic common trends) is consistent with the predictions of the NK-DSGE model, using a finite order VAR model. Next, we test the implied overidentifying cointegrating restrictions, conditional on the chosen rank. Finally, we test the CER the NK-DSGE model places on the VAR system, conditional on the cointegrating restrictions. Overall, the suggested method involves computing a sequence of LR tests, called $LR1$ (LR cointegration rank test), $LR2$ (LR cointegration matrix test) and $LR3$ (LR test for CER). The test $LR2$ is run conditional upon $LR1$ not rejecting the cointegration rank and $LR3$ is run if $LR2$ does not reject the overidentifying cointegration restrictions. To our knowledge, King et al. (1991), Canova et al. (1994) and Söderlind and Vredin (1996) are early examples of the use of the test $LR1$ in related contexts. Juselius (2011) is a recent example of the use of the test $LR2$ in the NK-DSGE models, while Guerron-Quintana et al. (2013) propose the inversion of a test like $LR3$ to build confidence sets for structural parameters that will be robust to identification failure. For ease of exposition, we denote our testing strategy with the symbol '$LR1 \to LR2 \to LR3$'. The novelty of the '$LR1 \to LR2 \to LR3$' procedure is that the empirical evaluation of the NK-DSGE model is treated as a multiple hypothesis testing approach. This is one of the contributions of our approach, as we will show in the rest of the paper.

Under the null of the NK-DSGE model, the tests $LR1$, $LR2$ and $LR3$, individually considered, are correctly sized in the sense that their asymptotic size is equal to the pre-fixed nominal type I error. Accordingly, using simple Bonferroni arguments, we can prove that the overall asymptotic size of the testing strategy does not exceed the sum of the type I errors pre-fixed for each test. If a practitioner wishes to test the NK-DSGE model at, say, the 5% nominal level of significance, the critical values of the tests $LR1$, $LR2$ and $LR3$ can be chosen such that the sum of the individual type I errors is 5%. The size of the overall testing strategy can be kept under strict control in small samples by referring to the bootstrap analogue of the '$LR1 \rightarrow LR2 \rightarrow LR3$' procedure. In this case, the bootstrap version of the test $LR1$ is computed following Cavaliere et al. (2012), and the bootstrap counterpart of the test $LR2$ is computed as in Boswijk et al. (2013), while the bootstrap analogue of the test $LR3$ can be computed as in, e.g., Cho and Moreno (2006) or Fanelli and Palomba (2011).

A discrepancy is often found between what the data tell and what theory implies when long-run restrictions are tested in structural forward-looking models. For instance, the balanced-growth-path property of the standard neoclassical growth model implies that hours worked are stationary. This, however, appears to be at odds with the persistent movements of per capita hours in the data. Similarly, NK-DSGE models typically maintain that inflation is a stationary process. In small samples, however, we typically observe high inflation persistence. The possible failures of the common-trend/cointegration restrictions using the tests $LR1$ and $LR2$ are generally the hardest features to interpret because of the lack of indications about how to modify the model. Chang et al. (2007) illustrate how the specification of a real business cycle DSGE model can be modified to incorporate non-stationary labor supply shocks which generate permanent shifts in hours worked. Similarly, Juselius (2011) provides a detailed interpretation of monetary business cycle NK-DSGE models under different scenarios reflecting the common trends that might be found in the data. As we show below, proper modifications to the probabilistic structure of the exogenous shocks that generate fluctuations in NK-DSGE models can be used to generalize trend structures and close the gap between theory and data. Likewise, when the short-run CER implied by the NK-DSGE model are rejected by the $LR3$ test, one should think of alternative structural frameworks to capture the dynamic features of the data or the omitted transmission mechanisms of the shocks. For example, dynamically rich, distributed-lag small scale monetary models have been employed by, e.g., Estrella and Fuhrer (2002, 2003) and Fuhrer and Rudebusch (2004), among others, while medium-scale systems which involve a relatively larger number of variables are considered in Christiano et al. (2005) and Smets and Wouters (2007). Thus, we go beyond using the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy as an 'accept-reject' proposition. We show that the outcomes of the individual tests can be used constructively to uncover what features of the data are not captured by the theoretical model and to rectify, when possible, the specification of the NK-DSGE model.

We evaluate the empirical performance of the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy by a small Monte Carlo experiment whose data generating process (DGP) belongs to the monetary business cycle NK-DSGE model discussed in Benati and Surico (2009), which is the leading example used in our paper. We further show the empirical usefulness of our approach by estimating and testing the NK-DSGE model of Benati and Surico (2009) using U.S. quarterly data.

Our paper has several connections with the existing literature. Canova et al. (1994) and Söderlind and Vredin (1996) propose a method to evaluate real business cycle models by eliciting the (highly) restricted VAR representation underlying them and comparing it with an unrestricted VAR for the data. They recognize that the driving forces in these models may be integrated, and hence account for the implied set of cointegration restrictions, as well as considering what Canova et al. (1994) call 'non-cointegrating restrictions'. Our approach differs from Canova et al. (1994) and Söderlind and Vredin (1996) in the way the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy is designed. Fukač and Pagan (2010) propose an evaluation approach to NK-DSGE models in which both the long and short-run behavior of the data are taken into account by modelling the common stochastic trends in an error-correction framework. While Fukač and Pagan (2010) put forth a 'limited information' approach, our analysis is developed in a 'full information' maximum likelihood (ML) framework. Also, Juselius (2011) applies a 'full information' ML approach, but he limits his attention to the steady-state implications of NK-DSGE models, leaving the CER untested. Gorodnichenko and Ng (2010) propose robust estimators for the parameters of DSGE models that do not require researchers to take a stand on whether shocks have permanent or transitory effects, while 'filtering' is implicitly obtained in our framework by a proper transformation of the model through the cointegration restrictions. The approach of Gorodnichenko and Ng (2010) is therefore suitable when the exact underlying cointegrating relationships are not known. Moreover, Gorodnichenko and Ng (2010) are not concerned with assessing how far/close is the estimated model from/to the data. One advantage of our method is that if the NK-DSGE model is not rejected by the data, it automatically delivers the ML estimates of the structural parameters, while the extension of Gorodnichenko and Ng's (2010) method to the case of ML estimation is not always practical.

Finally, apart from the Bayesian approach, we have many points in common with the 'DSGE-VAR' approach of Del Negro et al. (2007). These authors also use a cointegrated VAR in error-correction form as the statistical model for the data, but they impose the common-trend restrictions without testing. The prior distribution for the VAR parameters in Del Negro et al. (2007) is centered on the CER implied by the DSGE model and has dispersion governed by a scalar (hyper)parameter, denoted λ, such that small values of λ indicate that the VAR is far from the theoretical model, while large values of λ indicate that the theoretical model is supported by the data. A cutoff value for λ is not provided, as noticed by Christiano (2007). In our testing strategy, the test $LR3$ plays a role similar to λ in Del Negro et al. (2007). However, we have by construction a cut-off value for $LR3$ which depends on pre-fixed nominal type-I error: values of $LR3$ smaller than the cutoff value indicate that the VAR is 'close' to the NK-DSGE model, and vice versa.

The paper is organized as follows. We introduce the baseline NK-DSGE model and its assumptions in Section 2 and discuss a set of testable restrictions, which are usually ignored in the literature, in Section 3. We present our testing strategy in Section 4 and investigate its empirical size performance by a simulation experiment in Section 5. We present an empirical illustration in which our reference NK-DSGE model is evaluated on U.S. quarterly data in Section 6. Section 7 concludes the paper. The Appendix discusses the asymptotic size of the testing strategy. Additional details about the ML estimation algorithm for the structural parameters necessary to compute the test $LR3$, as well as the interpretation of the estimates obtained in the simulation experiment and in the empirical illustration, are reported in a Technical Supplement; see Bårdsen and Fanelli (2013).

2 Model and assumptions

Our starting point is the structural representation of a typical NK-DSGE model that aims at capturing the stylized features of the business cycle. The model is in the form of a system resulting from the log-linearization around steady-state values of the equations that describe the behavior of economic agents.

Let W_t be the p-dimensional vector collecting all the variables of the model of interest. A typical structural monetary NK-DSGE model takes the form of the linearized rational expectations model:

$$A_0^W W_t = A_f^W E_t W_{t+1} + A_b^W W_{t-1} + \eta_t^W, \tag{1}$$

where A_0^W, A_f^W and A_b^W are $p \times p$ matrices whose elements depend on the structural parameters collected in the vector θ, and η_t^W is a mean zero vector of structural disturbances. The term $E_t W_{t+1} = E(W_t \mid \mathcal{F}_t)$ denotes conditional expectations, where $\mathcal{F}_t$ is the available stochastic information set at time t and is such that $\sigma(W_t, W_{t-1},..., W_1) \subseteq \mathcal{F}_t$, and $\sigma(W_t, W_{t-1},..., W_1)$ is the sigma field generated by the variables. As is standard in the literature, we posit that the structural disturbance term η_t^W obeys a vector autoregressive processes of order one, i.e.,

$$\eta_t^W = R_W \ \eta_{t-1}^W + u_t^W \quad , \quad u_t^W \sim \mathrm{WN}(0_{p\times 1}, \Sigma_{W,u}) \tag{2}$$

where R_W is a diagonal stable matrix (i.e. with eigenvalues lying inside the unit disk) and u_t^W is a White Noise term with covariance matrix $\Sigma_{W,u}$. Hereafter, u_t^W will be denoted the vector of structural or 'fundamental' shocks, and it will be assumed that $\dim(u_t^W)$=$\dim(W_t) = p$, preventing the occurrence of the 'stochastic singularity' issue; see e.g. Ireland (2004) and DeJong and Dave (2011). In general, theory does not provide information about the correlation of the structural disturbances across equations; if cross-equation correlations are assumed for the structural disturbances, these can be captured by specifying a non-diagonal covariance matrix $\Sigma_{W,u}$. In our setup, the non-zero elements of R_W and of $vech(\Sigma_{W,u})$ belong to the vector of structural parameters θ. All meaningful values of θ belong to the 'theoretically admissible' (compact) parameter space, denoted $\mathcal{P}$.

A solution of model (1)-(2) is any stochastic process $\{W_t^*\}_{t=0}^{\infty}$, $W_t^* = W_t^*(\theta)$, such that, for $\theta \in \mathcal{P}$, $E_t W_{t+1}^* = E(W_{t+1}^* \mid \mathcal{F}_t)$ exists and if W_t^*, is substituted for W_t into the structural equations, and 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 W_t depends on u_t^W, lags of W_t and u_t^W (and, possibly, other arbitrary martingale difference sequences (MDS) with respect to $\mathcal{F}_t$ independent of u_t^W, called 'sunspot shocks').

We confine the class of reduced-form solutions associated with the NK-DSGE model to a known family of linear models by the assumption that follows.

Assumption 1 [Determinacy] The 'true' value θ_0 of θ is an interior point of $\mathcal{P}^*$, where $\mathcal{P}^* \subset \mathcal{P}$ is such that, for each $\theta \in \mathcal{P}^*$, the NK-DSGE model (1)-(2) has a unique and asymptotically stationary (stable) reduced-form solution.

Assumption 1 can be interpreted as the null hypothesis that the DGP belongs to the unique stable solution of system (1)-(2). This assumption is standard in the literature on NK-DSGE models and hinges on the idea that the time series upon which model (1) is built and estimated are typically constructed (or conceptualized) as stationary deviations from steady-state values.

Under Assumption 1, the unique stable solution of the model (1)-(2) can be represented as the asymptotically stationary VAR system

$$W_t = \tilde{F}_1 W_{t-1} + \tilde{F}_2 W_{t-2} + \varepsilon_t^W \quad , \quad \varepsilon_t^W = \tilde{Q} u_t^W \tag{3}$$

where $\tilde{F}_1 = F_1(\theta)$, $\tilde{F}_2 = F_2(\theta)$ and $\tilde{Q} = Q(\theta)$ are $p \times p$ matrices that depend nonlinearly on θ through the implicit set of nonlinear CER:

$$(A_0^{WR} - A_f^W \tilde{F}_1)\tilde{F}_1 - A_f^W(\tilde{F}_2) + A_{b,1}^W = 0_{p\times p} \tag{4}$$

$$(A_0^{WR} - A_f^W \tilde{F}_1)\tilde{F}_2 - A_{b,2}^W = 0_{p\times p} \tag{5}$$

$$\tilde{\Sigma}_{W,\varepsilon} = \tilde{Q}\ \Sigma_{W,u}\ \tilde{Q}'. \tag{6}$$

In Eqs. (4)-(6), $A_0^{WR} = (A_0^W + R_W A_f^W)$, $A_{b,1}^W = (A_b^W + R_W A_0^W)$, $A_{b,2}^W = -R_W A_b^W$, $\tilde{Q} = Q(\theta) = \left(A_0^W - A_f^W \tilde{F}_1\right)^{-1}$, and $\tilde{\Sigma}_{W,\varepsilon}$ is the constrained covariance matrix of ε_t^W, see Binder and Pesaran (1995), Uhlig (1999), Kapetanios et al. (2007) and Fanelli (2012).

In general, the NK-DSGE model represented in Eq.s (1)-(2) reads as a 'partial equilibrium' model, in the sense that it does not specify how any unobservable components of W_t, denoted W_t^u, are generated. For instance, in the examples we discuss below, the NK-DSGE model specified in the form (1)-(2) is based on the output gap and takes as given the process generating the natural level of output.

Let W_t^o be the sub-vector of W_t that contains the observable variables. Given the n-dimensional 'complete' vector $Z_t = (W_t^{o\prime}, W_t^u)'$, $n \geq p$, which collects, without any loss of generality, the observed (first) and unobserved variables (last), one can interpret the (stationary) vector W_t in systems (1) and (3) as obtained from the linear combination

$$W_t{=}\zeta' Z_t \tag{7}$$

where ζ is a known $n \times p$ matrix of full column-rank p that combines the observed and unobserved variables and/or picks out the stationary elements of Z_t that enter the structural model. A further step toward the 'complete' specification of the NK-DSGE model is provided by Assumption 2.

Assumption 2 [Unobserved processes are integrated of order one] The sub-vector W_t^u is such that ΔW_t^u is covariance stationary.

Assumption 2 simply states that W_t^u is integrated of order one, denoted $W_t^u \sim I(1)$. It does not provide a detailed specification of the process generating the unobserved variables, but it can be further specialized according to the specific features of the model under investigation, as shown in the next sub-sections. Given the scope of the present paper, approximating the unobserved components with $I(1)$ processes meets two requirements. First, Assumption 2 formalizes that the NK-DSGE model features stochastic trends. Second, the $I(1)$ assumption may represent a reasonable and interpretable choice for the typical unobservable components which characterize the class of small-scale NK-DSGE models used in monetary policy and business cycle analysis, i.e., the natural level of output (potential output) and/or the inflation target (or trend inflation), as suggested by Bekaert et al. (2010), Fukač and Pagan (2010) and Section 2.1 below.

Given Assumptions 1-2 and a detailed specification of the process generating W_t^u, the 'complete' (fully-specified) NK-DSGE model can be given the structural representation

$$A_0^Z Z_t = A_f^Z E_t Z_{t+1} + A_b^Z Z_{t-1} + \eta_t^Z \tag{8}$$

$$\eta_t^Z = R_Z \eta_{t-1}^Z + u_t^Z \quad , \quad u_t^Z \sim \text{WN}(0_{n\times 1}, \Sigma_{u,Z}), \tag{9}$$

where the matrices A_0^Z, A_f^Z, A_b^Z and $\Sigma_{u,Z}$ not only depend on θ, but also on a set of additional parameters, denoted with θ^a, that are associated with the processes specified for W_t^u. The 'extended' vector of structural parameters is therefore given by $\theta^e = (\theta', \theta^{a\prime})'$. Compared to the formulation (1)-(2) of the NK-DSGE model, the system represented in Eqs. (8)-(9) incorporates the unit-root implication of Assumption 2. We will refer to the representation in Eqs. (8)-(9) as the 'complete' representation of the NK-DSGE model. It is worth remarking that, albeit W_t in Eq. (7) reads as a sub-vector of Z_t, W_t has the finite-order VAR representation in Eq.s (3)-(6) under Assumption 1.

The next sub-section provides a detailed example about the relationship between the representation in Eq.s (1)-(2) and (8)-(9) of the NK-DSGE model.

2.1 An example model

We use an example based on Benati and Surico (2009). The model consists of the following equations:

$$\tilde{y}_t = \gamma E_t \tilde{y}_{t+1} + (1-\gamma)\tilde{y}_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \eta_{\tilde{y},t} \tag{10}$$

$$\pi_t = \frac{\varrho}{1+\varrho\varkappa} E_t \pi_{t+1} + \frac{\varkappa}{1+\varrho\varkappa} \pi_{t-1} + \kappa \tilde{y}_t + \eta_{\pi,t} \tag{11}$$

$$i_t = \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_y \tilde{y}_t) + \eta_{i,t} \tag{12}$$

$$\eta_{a,t} = \rho_a \eta_{a,t-1} + u_{a,t} \quad , \quad u_{a,t} \sim WN\left(0, \sigma_a^2\right) \ , \ a = \tilde{y}, \pi, i \tag{13}$$

hence $W_t = (\tilde{y}_t, \pi_t, i_t)'$, $p = 3$, $\eta_t^W = (\eta_{\tilde{y},t}, \eta_{\pi,t}, \eta_{i,t})'$ and $u_t^W = (u_{\tilde{y},t}, u_{\pi,t}, u_{i,t})'$. In this model, $\tilde{y}_t = (y_t - y_t^p)$ is the output gap, where y_t is the log of output and y_t^p the natural rate of output; π_t is the inflation rate and i_t is the nominal interest rate; $\eta_{\tilde{y},t}$, $\eta_{\pi,t}$ and $\eta_{i,t}$ are stochastic disturbances autocorrelated of order one, while

$u_{\tilde{y},t}$, $u_{\pi,t}$ and $u_{i,t}$ can be interpreted as demand, supply and monetary shocks, respectively. The structural parameters are collected in the vector $\theta = (\gamma,\, \delta,\, \varrho,\, \varkappa,\, \kappa, \rho,\, \varphi_\pi,\, \varphi_y,\, \rho_{\tilde{y}},\, \rho_\pi,\, \rho_i,\, \sigma^2_{\tilde{y}}, \sigma^2_\pi, \sigma^2_i)'$ and their economic interpretation may be found in Benati and Surico (2009).

The model (1)-(2) is 'incomplete' according to our definition, because it does not specify the process for the natural level of output y^p_t. One way to 'complete' the model is to specialize Assumption 2 as follows:

Assumption 2' [Potential output is a Random Walk]

$$y^p_t = y^p_{t-1} + \eta_{y^p,t} \quad , \quad \eta_{y^p,t} \sim \mathrm{WN}(0, \sigma^2_{y^p}). \tag{14}$$

In addition to the non-stationarity hypothesis, Assumption 2' provides a simple DGP for y^p_t consistent with the representation in Eq.s (8)-(9), see below. The usual interpretation of Assumption 2' is that the flexible price level of output y^p_t is driven by a combination of a stationary demand shock and a non-stationary technology shock, as in Ireland (2004). In this framework, the vector $W_t = (\tilde{y}_t, \pi_t, i_t)'$ can be thought of as being obtained through the linear combination in Eq. (7), which here qualifies in the expression

$$W_t = \underset{\zeta'}{\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}} \underset{Z_t}{\begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y^p_t \end{pmatrix}}. \tag{15}$$

The vector $Z_t = (W^{o\prime}_t, W^{u\prime}_t)'$ accommodates both the observed $W^o_t = Z^o_t = (y_t, \pi_t, i_t)'$ and the unobserved $W^u_t = y^p_t$ variables. Given the relationship in Eq. (15), the Eq.s (10)-(13) jointly with Eq. (14) imply the following configuration of the matrices A_0, A_f and A_b:

$$A^Z_0 = \begin{pmatrix} 1 & 0 & \delta & -1 \\ -\kappa & 1 & 0 & \kappa \\ -(1-\rho)\,\varphi_y & -(1-\rho)\,\varphi_\pi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A^Z_f = \begin{pmatrix} \gamma & \delta & 0 & -\gamma \\ 0 & \omega_f & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^Z_b = \begin{pmatrix} (1-\gamma) & 0 & 0 & -(1-\gamma) \\ 0 & \omega_b & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $\omega_f = \varrho/(1+\varrho\varkappa)$ and $\omega_b = \varkappa/(1+\varrho\varkappa)$. The 'extended' vector of parameters is $\theta^e = (\theta', \theta^a)'$, where $\theta^a = \sigma^2_{y^p}$.

3 Testable restrictions

The relationships between W_t and Z_t defined by Eq. (7) and the representation in Eq.s (8)-(9) of the NK-DSGE model can conveniently be used to analyze the whole set of testable restrictions at low and high frequencies. Under Assumptions 1-2, $Z_t \sim I(1)$, and all cointegration/common-trend restrictions of the system are subsumed in the vector Z_t. The model which captures the CER after factoring out the cointegrating relations from the system is given by the finite-order VAR representation for W_t in Eq.s (3)-(6). In this section, we explore the set of testable implications related to the vector Z_t, while in the next section we exploit these implications of the NK-DSGE model to define a coherent testing strategy.

We consider the n-dimensional vector of transformed variables

$$Y_t = \begin{pmatrix} \beta_0' \\ \tau'(1-L) \end{pmatrix} Z_t = G(\beta_0, \tau, 1-L) Z_t, \quad det(\tau' \beta_{0\perp}) \neq 0, \tag{16}$$

where β_0 is the $n \times r$ identified cointegration matrix, and τ is a $(n-r) \times r$ selection matrix, which is restricted so that it is not orthogonal to $\beta_{0\perp}$. The role of τ is to pick out a proper set of variables in first differences from the vector $(1-L) Z_t = \Delta Z_t$, where L is the lag operator ($L^j Z_t = Z_{t-j}$). The choice of τ in Eq. (16) is not necessarily unique, however. The case discussed below shows that, despite the many possible choices of τ, not all of them are consistent with the theoretical features of the NK-DSGE model. In principle, β_0 may temporarily depend on some 'additional' parameters that we collect in the vector ν, and which are not necessarily related to θ. We write $\beta_0 = \beta_0(\nu)$ to make clear such a dependence. Under the null hypothesis that the NK-DSGE model is valid, and with all constraints implied by the NK-DSGE model imposed on the system for Z_t, the joint restriction

$$r = p \quad , \quad \beta_0 = \beta_0^b = \zeta \tag{17}$$

must hold, where the symbol β_0^b denotes the counterpart of the identified cointegration matrix β_0 that leads to what we shall define below as a 'balanced' (error-correction) representation of the NK-DSGE model. Eq. (17) maintains that, under the null hypothesis that the NK-DSGE model is 'true', the identified cointegration matrix β_0^b must be equal to the selection matrix ζ of Eq. (7) and, accordingly, must not depend on any parameter. Hence, the dependence of β_0 on ν is suppressed in Eq. (17). We observe that the transformation in Eq. (16) mimics the one used by Campbell and Shiller (1987) to address the analysis of present value models.

Under the restriction (17), we can recover W_t from Y_t as follows:

$$
\begin{aligned}
Y_t = & \begin{pmatrix} \beta_0' \\ \tau'(1-L) \end{pmatrix} Z_t = G(\beta_0^b, \tau, 1-L) Z_t \\
= & \begin{pmatrix} \zeta' \\ \tau'(1-L) \end{pmatrix} Z_t = \begin{pmatrix} \zeta' Z_t \\ \tau' \Delta Z_t \end{pmatrix} = \begin{pmatrix} W_t \\ \tau' \Delta Z_t \end{pmatrix}
\end{aligned}
$$

hence it is seen that the vector W_t is part of Y_t. Because the $G(\beta_0, \tau, 1-L)$ (or $G(\beta_0^b, \tau, 1-L)$) matrix in system (16) is non-singular by construction, the mapping in Eq. (16) can be used in the model (8) to obtain

$$
A_0^Z G(\beta_0, \tau, 1-L)^{-1} Y_t = A_f^Z G(\beta_0, \tau, 1-L)^{-1} E_t Y_{t+1} + A_b^Z G(\beta_0, \tau, 1-L)^{-1} Y_{t-1} + \eta_t^Z . \tag{18}
$$

The appealing feature of the representation in Eq. (18) is that, other than involving stationary variables (i.e. those in Y_t), the (inverse of the) difference operator $(1-L)$ cancels out from the equations if one restricts β_0, as in Eq. (17), and imposes a proper set of restrictions on θ, such that the transformed model is 'balanced'. With the term 'balanced,' we mean that $G(\beta_0, \tau, 1-L)^{-1}$ is replaced with $G(\beta_0^b, \tau, 1-L)^{-1}$ and some restrictions are placed on the elements of θ, so all left-hand and right-hand side variables appearing in system (17) variables are stationary. The nature of these restrictions will be demonstrated in the examples that follow.

Hereafter we use the representation

$$
A_0^Y Y_t = A_f^Y E_t Y_{t+1} + A_b^Y Y_{t-1} + \eta_t^Y \tag{19}
$$

$$
\eta_t^Y = R_Y \eta_{t-1}^Y + u_t^Y \tag{20}
$$

to denote the 'balanced' counterpart of system (18). The system (19)-(20) can be regarded as an error-correction representation of the NK-DSGE model.

The structural parameters in the matrices A_0^Y, A_f^Y, A_b^Y, R^Y and $\Sigma_{Y,u} = E\left(u_t^Y u_t^{Y\prime}\right)$ are collected in the vector θ^Y, where θ^Y is obtained from θ^e by imposing the restrictions that map system (18) into the transformed representation in (19)-(20). In general, $\dim(\theta^Y) = \dim(\theta^e) - c$, where c is the total number of restrictions on θ^e necessary for balancing. Under Assumptions 1-2 (and the other minor assumptions in Section 2), if the unique stable solution of the NK-DSGE model (19)-(20) exists, it can be represented in the form

$$Y_t = \tilde{\Phi}_1 Y_{t-1} + \tilde{\Phi}_2 Y_{t-2} + \varepsilon_t^Y \quad , \ \varepsilon_t^Y = \tilde{\Psi} u_t^Y \tag{21}$$

where $\tilde{\Phi}_1 = \Phi_1(\theta^Y)$, $\tilde{\Phi}_2 = \Phi_2(\theta^Y)$ and $\tilde{\Psi} = \Psi(\theta^Y) = (A_0^{Y,R} - A_f^Y \tilde{\Phi}_1)^{-1}$ are $n \times n$ matrices that depend nonlinearly on θ^Y through the set of nonlinear CER:

$$(A_0^{Y,R} - A_f^Y \tilde{\Phi}_1)\tilde{\Phi}_1 - A_f^Y \tilde{\Phi}_2 + A_{b,1}^{Y,R} = 0_{n\times n} \tag{22}$$

$$(A_0^{Y,R} - A_f^Y \tilde{\Phi})\tilde{\Phi}_2 - A_{b,2}^{Y,R} = 0_{n\times n} \tag{23}$$

$$\tilde{\Sigma}_{Y,\varepsilon} = \tilde{\Psi} \ \Sigma_{Y,u} \tilde{\Psi}' \tag{24}$$

where $A_0^{Y,R} = (A_0^Y + R_Y A_f^Y)$, $A_{b,1}^{Y,R} = (A_b^Y + R_Y A_0^Y)$, $A_{b,2}^{Y,R} = -R_Y A_b^Y$, and $\tilde{\Sigma}_{Y,\varepsilon}$ is the covariance matrix of the reduced form disturbances ε_t^Y under the constraints, see Section 2. The constraints in Eq.s (22)-(24) mimic those derived in Eq.s (4)-(6) for the 'original' specification of the NK-DSGE model, but here refer to the 'complete' specification based on Assumption 2.

We now come back to our leading example, analyzed in Sub-section 2.1, to discuss some situations which clarify the essence of the transformations from Z_t to Y_t and the resulting set of testable restrictions. In particular, Sub-section 3.1 deals with the 'desirable' situation in which the number of common stochastic trends found in the data, $n - r$, lines up with the number predicted by the theory, $n - p$, and the identified cointegration relationships match perfectly the structure in Eq. (17). Sub-section 3.2 deals, instead, with the situation where the inferred number of common trends, $n - r$, is larger than $n - p$ ($r < p$), hence the structure of the cointegration relationships in Eq. (17) is no longer valid. We argue that, in these cases, it is generally possible to rectify the specification of the structural equations to feature the 'additional' stochastic trends, provided these trends can be given a sensible economic interpretation. This process, however, may give rise to extra (testable) restrictions in other parts of the model. Finally, Sub-section 3.3 deals with the case in which it is necessary to think about alternative specifications of the transmission mechanisms of the shocks. This situation may arise when the CER in Eq.s (22)-(24) (or in Eq.s (4)-(6) if the 'complete' specification is investigated) are not valid.

3.1 Example 1: the case of a single stochastic trend

Consider the NK-DSGE model in Eqs. (10)-(13) and Assumption 2'. Suppose that given $Z_t = (y_t, \pi_t, i_t, y_t^p)'$, the selected cointegration rank r and the identified cointegration relationships β_0 perfectly match the requirements in Eq. (17), i.e.

$$r = 3, \ \ \beta_0' Z_t = \beta_0^{b\prime} Z_t = \zeta' Z_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = W_t. \tag{25}$$

In this case, the output gap, inflation and the short term interest rate are jointly stationary and there is a single stochastic trend in Z_t. The cointegration relationships in Eq. (25) are consistent with the hypothesis that the system is driven by a non-stationary technology shock; see, e.g., Ireland (2004) and DeJong and Dave (2011).

The vector Y_t in Eq. (16) is given by

$$Y_t = G(\beta_0^b, \tau, 1-L) Z_t = \begin{pmatrix} \beta_0^{b\prime} \\ \tau'(1-L) \end{pmatrix} Z_t = \begin{pmatrix} & \beta_0^{b\prime} & & \\ (1-L) & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = (W_t', \Delta y_t)', \tag{26}$$

where it can be noticed that $\tau = (1,0,0,0)'$, $\beta_{0\perp} = (1,0,0,1)'$, hence $\det(\tau'\beta_{0\perp}) = \det(1) \neq 0$. Using $G(\beta_0^b, \tau, 1-L)^{-1}$ in Eq. (18) generates the system

$$\begin{aligned} \tilde{y}_t &= \gamma E_t \tilde{y}_{t+1} - \delta(i_t - E_t \pi_{t+1}) + (1-\gamma)\tilde{y}_{t-1} + \eta_{\tilde{y},t} \\ \pi_t &= \omega_f E_t \pi_{t+1} + \omega_b \pi_{t-1} + \kappa \tilde{y}_t + \eta_{\pi,t} \\ i_t &= \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_y \tilde{y}_t) + \eta_{i,t} \\ -\tilde{y}_t + (1-L)^{-1}(1-L) y_t &= -\tilde{y}_{t-1} + (1-L)^{-1}(1-L) y_{t-1} + \eta_{y^p,t}, \end{aligned}$$

where we have left the operator $(1-L)^{-1}$ in the final equation to highlight the point about balancing. To see that $(1-L)^{-1}$ cancels out from the former equation, it is sufficient to rewrite it in the form (using Assumption 2')

$$\tilde{y}_t = \tilde{y}_{t-1} + \Delta y_t + \eta^*_{y^p,t} \tag{27}$$

where $\eta^*_{y^p,t} = -\eta_{y^p,t}$. It is worth emphasizing that Eq. (27) is a reparametrization of the Random Walk model assumed for y_t^p in Eq. (14). A similar representation obtains if the chosen τ in $G(\beta_0^b, \tau, 1-L)$ is given by $\tau = (0,0,0,1)'$.

In this case, $\theta^Y = \theta^e$ and the matrices A_0^Y, A_f^Y and A_b^Y, as well as the vector η_t^Y in the representation in Eq. (19), can easily be derived and are equal to

$$A_0^Y = \begin{pmatrix} 1 & 0 & \delta & 0 \\ -\kappa & 1 & 0 & 0 \\ -(1-\rho)\varphi_y & -(1-\rho)\varphi_\pi & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \quad A_f^Y = \begin{pmatrix} \gamma & -\delta & 0 & 0 \\ 0 & \omega_f & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A_b^Y = \begin{pmatrix} 1-\gamma & 0 & 0 & 0 \\ 0 & \omega_b & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

and $\eta_t^Y = (\eta_{\tilde{y},t}, \eta_{\pi,t}, \eta_{i,t}, \eta^*_{y^p,t})'$. The testable cointegration restrictions relative to the strictly observable time series in Z_t, $Z_t^o = (y_t, \pi_t, i_t)' = W_t^o \sim I(1)$ are

$$r^o = 2 \quad , \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} Z_t^o \sim \text{stationary} \tag{28}$$

where r^o is the cointegration rank associated with Z_t^o .

3.2 Example 2: the case of two stochastic trends

Consider the NK-DSGE model in Eqs. (10)-(14). Imagine that, given $Z_t = (y_t, \pi_t, i_t, y_t^p)'$, the selected cointegration rank is $r = 2 \neq p$, implying the existence of $n - r = 4 - 2 = 2$ stochastic trends, one more than the technology trend predicted by the baseline version of the model. We discuss in detail two possible scenarios which can be used to nest a setup like this, denoted Hypothesis 1 and Hypothesis 2, respectively.

Hypothesis 1: stochastic inflation target

Inflation is typically a highly persistent process, which can sometimes be approximated reasonably well by $I(1)$ processes. If the quantity $\frac{\varrho}{1+\varrho\varkappa} + \frac{\varkappa}{1+\varrho\varkappa}$ in the NKPC in Eq. (11) is close to 1, the $I(1)$ approximation for π_t is sensible in small samples. While it is implicitly assumed that trend inflation is zero in the system (10)-(13), it may be the case that trend inflation is determined by the long-run target of the central bank's policy rule. A drift in trend inflation could therefore be attributed to shifts in that target. Consider, for instance, the small monetary NK-DSGE model investigated by Bekaert et al. (2010). Their model differs from our leading example only in the specification of the policy rule, which in their framework is given by

$$i_t = \rho i_{t-1} + (1-\rho)\varphi_\pi(E_t\pi_{t+1} - \pi_t^*) + (1-\rho)\varphi_y\tilde{y}_t + \eta_{i,t} \tag{29}$$

where π_t^* is a stochastic inflation target generated by the equation

$$\pi_t^* = \frac{\varrho}{1+\varrho\varpi} E_t \pi_{t+1}^* + \frac{\varpi}{1+\varrho\varpi} \pi_{t-1}^* + (1 - \frac{\varrho}{1+\varrho\varpi} - \frac{\varpi}{1+\varrho\varpi})\pi_t + \epsilon_{\pi^*,t}. \tag{30}$$

$\epsilon_{\pi^*,t}$ in Eq. (30) is an exogenous shift in the policy stance regarding the long term rate of inflation, assumed to be i.i.d., and the parameter ϖ measures the extent to which the monetary authority smoothes the inflation target in anchoring its inflation target to a properly defined measure of the 'long-run inflation expectations', $\pi^{LR} = (1-\varrho)\sum_{j=0}^{\infty} \varrho^j E_t \pi_{t+j}$, see Eq. (9) in Bekaert et al. (2010). Therefore, with $\varpi = 0$, the target π_t^* equals long-run inflation expectations in the absence of exogenous shifts, while, for values of ϖ close or equal to unity, Eq. (30) collapses to the Random Walk model: $\pi_t^* = \pi_{t-1}^* + \epsilon_{\pi^*,t}$. If this model can be taken as a reasonable description of the evolution of π_t^* from the modelled sample, π_t and π_t^* must be cointegrated with cointegration vector (1,-1) for the monetary policy rule in Eq. (29) to be balanced (note that, if π_t-π_t^* is stationary, $(E_t\pi_{t+1} - \pi_t^*)$ also will be stationary). Under this DGP, the investigator will find two stochastic trends in $Z_t = (y_t, \pi_t, i_t, y_t^p)'$ and the source of the cointegration rank failure (i.e. $r < p$) will be the omitted stochastic inflation target. A necessary condition for this hypothesis to be valid is that it holds the restriction

$$\beta_0' Z_t = \beta_0^{b\prime} Z_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} y_t - y_t^p \\ i_t \end{pmatrix} \sim \text{stationary.} \tag{31}$$

The specification in Eq. (31) implies that y_t^p and π_t are $I(1)$ and $y_t - y_t^p$ and i_t are stationary. The testable cointegration restrictions relative to the strictly observable time series $Z_t^o = (y_t, \pi_t, i_t)' = W_t^o$ collapse to

$$r^o = 1 \quad , \quad (0,0,1,0)Z_t^o \sim \text{stationary.} \tag{32}$$

Hypothesis 2: stationary real (ex-post) interest rate

Suppose that the two cointegration relationships identified from the data are given by

$$\beta_0' Z_t = \beta_0'(\nu) Z_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -\nu & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} y_t - y_t^p \\ i_t - \nu\pi_t \end{pmatrix} \sim \text{stationary} \tag{33}$$

where ν is a cointegration parameter not related to the structural parameters θ^e. The two cointegrating vectors in Eq. (33) are the output gap and a linear combination of i_t and π_t with cointegration vector $(1, -\nu)$. For values of the parameter ν close to one, the stationary linear combination $(i_t - \nu\pi_t)$ can be interpreted as a measure of the ex-post real interest rate. Hence, other than the technology trend, we can think about a real interest rate trend (or Fisher parity relationship). The mapping from Z_t to Y_t is in this case given by

$$Y_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -\nu & 1 & 0 \\ 0 & (1-L) & 0 & 0 \\ (1-L) & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} \tilde{y}_t \\ i_t - \nu\pi_t \\ (1-L)\pi_t \\ (1-L)y_t \end{pmatrix} \sim \text{stationary}$$

and by imposing this transformation on the system (10)-(13), the analogue of the system (18) is given by

$$\begin{aligned}
\tilde{y}_t &= -\delta(i_t - \nu\pi_t) - \left(\frac{\nu\delta}{1-L}\right)\Delta\pi_t + \gamma E_t\tilde{y}_{t+1} + \left(\frac{\delta}{1-L}\right)E_t\Delta\pi_{t+1} + (1-\gamma)\tilde{y}_{t-1} + \eta_{\tilde{y},t} \\
\left(\frac{1}{1-L}\right)\Delta\pi_t &= \frac{\varrho}{1+\varrho\varkappa}\left(\frac{1}{1-L}\right)\Delta E_t\pi_{t+1} + \frac{\varkappa}{1+\varrho\varkappa}\left(\frac{1}{1-L}\right)\Delta\pi_{t-1} + \kappa\tilde{y}_t + \eta_{\pi,t} \\
(i_t - \nu\pi_t) &= (1-\rho)\varphi_y\tilde{y}_t - \left(\frac{\nu - (1-\rho)\varphi_\pi}{1-L}\right)\Delta\pi_t + \frac{\nu\rho}{(1-L)}\Delta\pi_{t-1} + \rho(i_{t-1} - \nu\pi_{t-1}) + \eta_{i,t} \\
\tilde{y}_t &= \tilde{y}_{t-1} + \Delta y_t + \eta^*_{y^p,t}.
\end{aligned}$$

In order to make only the stationary variables in Y_t enter the model, we need to eliminate the operator $(1-L)^{-1}$ from the equations above. This is achieved by imposing the following additional parameter restrictions: $\nu = \varphi_\pi = 1$; $\varkappa = 1$, which, after some algebraic manipulations, lead to the equations

$$\tilde{y}_t = \gamma E_t\tilde{y}_{t+1} + (1-\gamma)\tilde{y}_{t-1} + \delta E_t\Delta\pi_{t+1} - \delta(i_t - \pi_t) + \eta_{\tilde{y},t} \tag{34}$$

$$\Delta\pi_t = \varrho E_t\Delta\pi_{t+1} + \left(\frac{\kappa}{1+\varrho}\right)\tilde{y}_t + (1+\varrho)\eta_{\pi,t} \tag{35}$$

$$(i_t - \pi_t) = (1-\rho)\varphi_y\tilde{y}_t - \rho\Delta\pi_t + \rho(i_{t-1} - \pi_{t-1}) + \eta_{i,t} \tag{36}$$

$$\Delta\tilde{y}_t = \Delta y_t + \eta^*_{y^p,t} \quad (\text{or } \Delta y_t^p = \eta_{y^p,t}). \tag{37}$$

This new system is the counterpart of the error-correction representation of the NK-DSGE system (19)-(20). It is balanced because it does not involve variables other those in Y_t. The vector of structural parameters is $\theta^Y = (\gamma, \delta, \varrho, \rho, \varphi_y, \rho_{\tilde{y}}, \rho_\pi, \rho_i, \sigma^2_{\tilde{y}}, \sigma^2_\pi, \sigma^2_i, \sigma^2_{y^p})'$, and is obtained from the 'complete' vector $\theta^e = (\theta', \theta^a)'$ $(\theta^a = \sigma^2_{y^p})$ by imposing the constraints $\nu = \varphi_\pi = 1$; $\varkappa = 1$. Note that $\dim(\theta^Y)$=$\dim(\theta^e)$-2, where 2 is the number of restrictions required for balancing. The interpretation of the transformed system is not trivial. The policy rule in Eq. (36) is now expressed such that the 'operational/implementation target' is the real interest rate, while the decision variables are the output gap and the change in inflation (the so-called 'acceleration rate'). The condition $\varphi_\pi = 1$ maintains that the long-run response of the Central Bank to inflation is equal to one, which stands in sharp contrasts with a widely shared view about the conquest of the U.S. inflation during the 'Great Moderation' era; see e.g. Clarida et al. (2000) and Lubik and Schorfheide (2004). The restriction $\varkappa = 1$ implies 'full indexation' and the Phillips curve in Eq. (35) is expressed as a 'purely forward-looking' model for $\Delta\pi_t$. In this case, the testable cointegration restrictions relative to the strictly observable time series $Z_t^o = (y_t, \pi_t, i_t)' = W_t^o$ correspond to

$$r^o = 1, \quad (0, -\nu, 1, 0)Z_t^o \sim \text{stationary, with } \nu = 1. \tag{38}$$

One can evaluate Hypothesis 1 versus Hypothesis 2 by testing whether, for $r = 2$ ($r^o = 1$), the cointegration relationship is better described by the structure in Eq. (31) (Eq. (32)), or the structure in Eq. (33) (Eq. (38)), using the testing strategy we introduce in the next section.

3.3 Example 3: the cost channel

Consider again the leading example model in Eq.s (10)-(14) but suppose now that the short-run CER implied by this system, summarized in Eq.s (4)-(6) (Eq.s (22)-(23)), are not valid for the sample period used to evaluate the model. One might conjecture, for instance, that this occurs because the baseline specification does not account for a cost-channel which might be at work in the economy. In short, a share of firms in the economy might need to borrow resources to pay workers' wages before the final goods market opens. According to this theory, see Christiano et al. (2005) and Ravenna and Walsh (2006), the policy rate also should enter firms' marginal costs as a proxy of the interest rate paid on their loans, leading to the system

$$\begin{aligned}
\tilde{y}_t &= \gamma E_t \tilde{y}_{t+1} - \delta(i_t - E_t\pi_{t+1}) + (1-\gamma)\tilde{y}_{t-1} + \eta_{\tilde{y},t} \\
\pi_t &= \omega_f E_t \pi_{t+1} + \omega_b \pi_{t-1} + \kappa \tilde{y}_t + \kappa_i \kappa i_t + \eta_{\pi,t} \\
i_t &= \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_y \tilde{y}_t) + \eta_{i,t} \\
\Delta\tilde{y}_t &= \Delta y_t + \eta^*_{y^p,t}.
\end{aligned}$$

In this model, the parameter $0 \leq \kappa_i < 1$ captures the share of firms acceding the financial markets. The implied set of CER obtained with $0 < \kappa_i < 1$ are different in the absence of a cost channel based on $\kappa_i = 0$.

4 The test sequence

We consider the NK-DSGE model introduced in Section 2, Assumptions 1-2, and focus on the following hypotheses

$$H_0\text{: the DGP belongs to system (21)-(24) ; } H_1\text{: the DGP does not belong to system (21)-(24)} \tag{39}$$

The null in Eq. (39) will also be referred to as the null with respect to the hypothesis that 'the NK-DSGE model is valid'. H_0 implicitly maintains that the cointegration/common-trend restrictions subsumed in Eq. (17) are fulfilled, hence it is possible to map Z_t into Y_t. Thus, any testing strategy for H_0 against H_1 is a conditional decision rule.

To simplify the exposition without altering the logic of our method, we assume temporarily that all variables in Z_t (and hence in Y_t) can be observed. We turn to the role of unobservables later. Consider the VAR model for Z_t :

$$Z_t = \sum_{j=1}^{\ell} P_j Z_{t-j} + \mu d_t + \xi_t \quad , \quad \xi_t \sim \text{WN}(0_{n\times 1}, \Sigma_\xi) \tag{40}$$

where P_j, $j = 1, ..., \ell$ are $n\times n$ matrices of parameters, ℓ is the lag order, d_t is a vector including deterministic variables (constant, linear trend dummies, etc.) with associated matrix of coefficients, μ, and ξ_t is a White Noise disturbance. Consider also the corresponding error-correction representation

$$\Delta Z_t = \alpha\beta' Z_{t-1} + \sum_{j=1}^{\ell-1} \Theta_j \Delta Z_{t-1} + \mu d_t + \xi_t \quad , \quad \xi_t \sim \text{WN}(0_{n\times 1}, \Sigma_\xi) \tag{41}$$

where $\alpha\beta' = (\sum_{j=1}^{\ell} P_j - I_n)$, α is the $n \times r$ matrix of adjustment coefficients, β is the $n \times r$ cointegration matrix and $\Theta_j = -\sum_{h=j+1}^{\ell} P_h$, $j = 1, ..., \ell-1$. Suppose that the VAR lag order, ℓ, is determined from the data and that the vector of deterministic components, d_t, is selected in accordance with the time series features observed for the variables in Z_t. Our procedure is based on the following steps:

$LR1$ **[Cointegration rank test]** We estimate the VAR system (40) and test for the hypothesis that the cointegration rank is $r = p$, corresponding to $n - p$ common stochastic trends driving the system, against the alternative $r = n$, corresponding to a stationary system. We suggest using either the 'one-shot' version of the LR Trace test (Johansen, 1996), henceforth denoted $LR1$, or its 'sequential' version, denoted $LR1_{\text{seq}}$. The $LR1_{\text{seq}}$ test involves starting with $r = 0$ (n stochastic trends), testing in turn the hypothesis 'the cointegration rank is r' against 'the VAR is stationary' for $r = 0, ..., n-1$, until, for a given value of $r = \hat{r}$, the asymptotic p-value associated with the test statistics exceeds the chosen significance level. The bootstrap versions of these tests discussed in Cavaliere et al. (2012) can be applied in small samples. If the hypothesis $r = p$ is not rejected, we consider the next step. If instead the hypothesis $r = p$ is rejected, we suggest proceeding by thinking about an alternative specification which embodies, when possible, the stochastic trends not featured by the original specification of the NK-DSGE model; see as an example the cases discussed in Sub-section 3.2.

$LR2$ [**Overidentification cointegration restrictions test**] Given $r = p$, we move to the vector error-correction counterpart of the cointegrated VAR in Eq. (41), and fix the (identified) cointegration matrix β at the structure implied by the theoretical model, i.e., $\beta = \beta_0^b = \zeta$ as in Eq. (17). Then we compute a LR test, henceforth denoted with $LR2$, for the implied set of over-identifying restrictions; see Johansen (1996). The bootstrap versions of the test $LR2$ discussed in, e.g., Boswijk et al. (2013) can be used to improve the small sample performance. If the $LR2$ test does not reject the over-identifying restrictions, we build the transformed vector Y_t in Eq. (16) by keeping $\beta = \beta_0^b = \zeta$ fixed at the non-rejected structure, and consider the next step. If instead the $LR2$ test rejects the restrictions, one can proceed similarly to the case in which the $LR1$ ($LR1_{\text{seq}}$) test rejects the predicted number of stochastic trends.

$LR3$ [**Test for CER**] We estimate the finite-order VAR system in Eq. (21) unrestrictedly, i.e., leaving the matrices Φ_1, Φ_2 and $\Sigma_{Y,\varepsilon}$ unconstrained, and imposing the CER in Eq.s (22)-(24), using the ML algorithm summarized in the Technical Supplement. We thus compute a LR test for the CER, henceforth denoted $LR3$, and obtain the ML estimate of the structural parameters. The $LR3$ test can also be constructed by referring to the 'partial equilibrium' representation of the NK-DSGE model for W_t, given by the system in Eq.s (3)-(6). When computationally feasible, bootstrap versions of the test LR3 discussed in, e.g., Cho and Moreno (2006) and Fanelli and Palomba (2011) can be used in small samples.

The '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy is a novel approach in the literature. From a statistical viewpoint, H_0 in Eq. (39) is rejected in favour of H_1 if one of the three tests rejects, while it is accepted if all three tests pass. However, the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy need not be applied mechanically as an 'accept-reject' proposition. The information, stemming from the tests $LR1$ ($LR1_{\text{seq}}$), $LR2$ and $LR3$, can potentially be used to uncover which features of the data or transmission mechanisms of the shocks are not captured by the theoretical model, so that the model may be improved; see, e.g., the scenarios discussed in Sub-sections 3.2-3.3. In principle, it is possible to follow the alternative strategy based on imposing (without testing) the restrictions in Eq. (17) on the VAR, testing the CER alone. While such a strategy is advantageous when the restrictions imposed are 'true', one of its perils is that insisting that a root very close to unity is a stationary root, for example, may lead to large size distortions and power losses in tests for the CER in rational expectations models; see Johansen (2006) and Li (2007).

The '$LR1 \to LR2 \to LR3$' testing strategy discussed so far maintains that the econometrician observes all components of Z_t (and hence of Y_t). When it is not possible to proxy all variables, the testing strategy can be adapted. In these cases, the testable cointegration/common-trend implications of the NK-DSGE model reflect on the sub-vector $Z_t^o = W_t^o$ of Z_t. For instance, considering the examples discussed in Subsections 3.1 and 3.2, the relevant log run testable restrictions are given in Eq. (28), Eq. (32) and Eq. (38), respectively. In this case, W_t^o has a state-space (VARMA-type) representation under the null H_0 in Eq. (39) (see the Technical Supplement), and a finite-order VAR for W_t^o can provide, with qualifications, a reasonable approximation of its actual time series properties. The procedure is based on the following testing steps:

$LR1$ **[Cointegration rank test: the case of unobservables]** We specify a VAR for W_t^o similar to system (40) with a relatively 'large' ℓ, and use the test(s) LR1 and/or LR1_{seq} to select the cointegration rank r^o.

$LR2$ **[Overidentification cointegration restrictions test: the case of unobservables]** We use the vector error-correction counterpart of the cointegrated VAR similar to system (41) and, fixing r^o, we compute the $LR2$ test by considering the over-identifying cointegration restrictions.

$LR3$ **[Test for CER: the case of unobservables]** We compute the $LR3$ test by evaluating the likelihood function associated with the 'minimal state-space representation' corresponding to the VAR for Y_t in Eq. (21) under the CER in Eq.s (22)-(24) (or to the VAR for W_t in Eq. (3) under the CER in Eq.s (4)-(6)) with the Kalman filter; see, e.g., Ruge-Murcia (2007) and Fukač and Pagan (2010). We refer to Komunjer and Ng (2011) and Guerron-Quintana et al. (2013) for practical examples where the 'minimal state-space representation' is derived from the set of observationally equivalent non-minimal state-space representations. An alternative to the test LR3 based on the indirect inference approach may be found in the Technical Supplement.

Under the null H_0 in Eq. (39), the asymptotic properties of each of the three tests comprising the 'LR1→LR2 →LR3' testing strategy are known. The asymptotic properties of the tests $LR1$ ($LR1_{\text{seq}}$) and $LR2$ may be found in Johansen (1996), while the asymptotic properties of their bootstrap counterparts are discussed in Cavaliere et al. (2012) and Boswijk et al. (2013). The asymptotic properties of the test $LR3$ are standard (including its bootstrap analogue) under standard regularity conditions. We have postponed a detailed derivation of the asymptotic size properties of the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy to the Appendix. Because the three tests are asymptotically correctly sized under the null, if the test for H_0 against H_1 in Eq. (39) is conducted by fixing the overall significance level at, e.g., the 5% level, the type-I errors (and hence the critical values) of the tests $LR1$ ($LR1_{\text{seq}}$), $LR2$ and $LR3$ must be chosen accordingly, e.g. 1% for $LR1$ ($LR1_{\text{seq}}$), 2% for $LR2$ and 2% for $LR3$.

It is worth spending a few words on the tests $LR1$ and $LR1_{\text{seq}}$. In our framework, the use of the 'one-shot' cointegration rank test $LR1$ reflects the idea that the number of common trends driving the variables are given and equal to $n - r_0 = n - p$ under the null of the NK-DSGE model, where r_0 is the true cointegration rank. By construction, the $LR1$ test rules out all alternatives in which the number of common trends is, e.g., larger than $n - p$, and in which its asymptotic distribution is unknown, if, for instance, $r_0 < p$. The $LR1$ test must be applied, therefore, when one is confident that there are no more than $n - p$ stochastic trends in the data. Instead, the $LR1_{\text{seq}}$ test has power asymptotically against the alternative of a number of stochastic trends different from $n - p$. Nevertheless, its finite sample performance may be poor, as our simulation experiment in the next section will document, and therefore it should be applied and interpreted with caution. When the empirical evaluation of the NK-DSGE model is based on the $W_t^o = Z_t^o$ sub-vector, it is also possible to apply a number of alternative tests to $LR1$ and $LR1_{\text{seq}}$. These are reviewed in, e.g., Lütkepohl and Claessen (1997), which do not require estimating a fully identified VARMA-type model; see also Stock and Watson (1988).

5 Simulation experiment

To evaluate the finite sample size performance of the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy, we conduct a small Monte Carlo experiment based on the determinate solution of the model summarized in Eq.s (10)-(14). We fix the discount rate ϱ at the value $\varrho = 0.99$, and consider the estimation of $\omega_f = \varrho/(1 + \varrho\varkappa)$ and derive that of $\varkappa$ indirectly. Hence the vector of 'free' structural parameters is given by $\theta = (\gamma,\, \delta,\, \omega_f,\, \rho,\, \varphi_\pi,\, \varphi_y,$ $\rho_{\tilde{y}},\, \rho_\pi,\, \rho_i,\, \sigma^2_{\tilde{y}}, \sigma^2_\pi, \sigma^2_i)'$ and the 'extended' vector is $\theta^e = (\theta',\, \theta^a)'$, where $\theta^a = \sigma^2_{y^p}$ ($\eta_{y^p,t} = u_{y^p,t}$). The vector of fundamental shocks $u_t^Z = (u_{\tilde{y},t}, u_{\pi,t}, u_{i,t}, u_{y^p,t})'$ is assumed White Noise with Gaussian distribution and diagonal covariance matrix $\Sigma_{u,Z}$. The parameter vector θ is calibrated to the empirical estimates of Benati and Surico (2009); see, in particular, the last column of their Table 1 ('After the Volcker stabilization'). The variance of the natural rate of output $\sigma^2_{y^p}$ is fixed at a value considered reasonable. Recall that, in this setup, $\theta^Y = \theta^e$. The calibrated values of θ^e are reported in the first column of Table 1, Panel 2. In line with the developments in the empirical illustration of Section 6, we assume that the econometrician can observe the natural rate of output, y_t^p.

[Table 1 about here.]

We generate time series of size T=100, 200 and 500, and compute the test sequence '$LR1 \rightarrow LR2 \rightarrow LR3$' $M = 5000$ times. The overall nominal level of significance is fixed at 5% ($\psi = 0.05$), and the nominal type-I errors of the three tests are 1% for the $LR1$ test ($\psi_1 = 0.01$), 2% for the test $LR2$ ($\psi_2 = 0.02$) and 2% for the test $LR3$ ($\psi_3 = 0.02$). The (asymptotic) critical values are chosen accordingly.

The results are summarized in Table 1, Panel 1. For samples of size $T = 100$ and 200 we also compute the (i.i.d.) bootstrap versions of the tests. The implementation of the bootstrap version of the $LR1$ test follows Cavaliere et al. (2012), while the (i.i.d.) bootstrap version of the $LR2$ test is discussed in Boswijk et al. (2013). The bootstrap version of the $LR3$ test is computed by adapting to the (i.i.d.) non-parametric setup the procedure discussed in Fanelli and Palomba (2011).

We first focus on the empirical size of the components of the test sequence. The null hypothesis of the test $LR1$ is a single stochastic trend in the system ($r = p = 3$) and its empirical size is reported in the first row of Panel 1, labeled '$LR1_{\psi_1=0.01}\,(r = 3)$'. We notice that, unexpectedly, this test is slightly under-sized in samples of size $T = 100$. One would expect over-rejection but the finite sample performance of the $LR1$ test may well depend on the structure of the short-run dynamics of the system, which in our setup is 'special', i.e., highly restricted by the CER. The bootstrap-corrected version of the test produces similar results. $LR2$ tests the over-identification restrictions on the cointegration matrix β_0 implied by Eq. (25) and is asymptotically chi-square distributed with 3 degrees of freedom under the null. Its empirical size is reported in the second row of Panel 1, labeled '$LR2_{\psi_2=0.02}\,(\beta_0 = \zeta \mid LR1)$'. The test tends to be over-sized. For a sample size of $T = 100$, the empirical size is 7.2% as opposed to the 2% nominal size. However, the bootstrap version of the test guarantees a good size coverage, bringing the rejection frequency down to 2.2% for both $T = 100$ and $T = 200$. Finally, the empirical size of the test $LR3$ for the CER is reported in the third row of Panel 1, labeled '$LR3_{\psi_3=0.02}\,(CER \mid LR2)$'. To compute this test, we maximized the likelihood of the VAR system (21) under the constraints in Eq.s (22)-(24), using the iterative ML algorithm discussed in the Technical Supplement. The ML estimates of θ^e are discussed below. Under the null, LR3 is asymptotically chi-square with 28 degrees of freedom, where 28 is the difference between the number of unrestricted parameters in the VAR $(32 + 10)$ and the structural parameters ($\dim(\theta^e) = 14$). The empirical size is reasonably good in this case although the bootstrap counterpart of the test is under-sized in samples of $T = 100$.

The overall empirical rejection frequency associated with the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy is summarized in the seventh row of Panel 1. It can be noticed that, considering asymptotic critical values, it ranges from 10.6% ($T = 100$), via 7.1% ($T = 200$) to 5% ($T = 500$), with a nominal level of 5%. The bootstrap version of the testing strategy ensures a strict size control in small samples.

Table 1, Panel 2, reports the Monte Carlo means of the structural parameters with the Monte Carlo standard errors in parentheses. The structural parameters are recovered with surprising precision, the only exceptions being, in samples of size T=100, the parameters of the policy rule φ_y and φ_π, although estimation precision increases with the sample size. This lack of precision, which is usually ascribed to 'weak identification' issues, is a common finding and source of misunderstandings in the literature. The discussion of these issues goes beyond the scope of the present paper. We suggest an interpretation in the Technical Supplement.

[Table 2 about here.]

As observed in Section 4, the first test of the testing strategy can also be the 'sequential' cointegration rank test, $LR1_{\text{seq}}$. The results in Table 2 summarize the marginal acceptance frequencies of the hypotheses $r = \hat{r}$, $\hat{r} = 0, 1, 2, 3, 4$, considering samples of size $T = 100$ and $T = 200$. We also include the acceptance frequencies corrected with the bootstrap version of the $LR1_{\text{seq}}$ test. We notice that, in samples of size $T = 200$, the LR1_{seq} test performs as expected, selecting the 'true' cointegration rank, $\hat{r} = r_0 = 3$, in 71.2% of the simulations. Instead, results are less clear-cut in samples of size $T = 100$. We notice that the 'wrong' cointegration ranks 1 and 2 are selected in around 90% of the simulations, compared to the 'true' cointegration rank in only 9.4% of the simulations. This phenomenon reflects a well-known small sample (power) issue of the sequential cointegration rank test, and, in this case, the bootstrap correction does not seem to keep the risk of a wrong choice under control. The results of our Monte Carlo experiment suggest using $LR1_{\text{seq}}$ with caution in small samples, especially in the absence of a clear alternative about the number of stochastic trends.

Keeping these results in mind, we next turn to an empirical application of our testing strategy.

6 An estimated NK-DSGE model of the U.S. economy

In this section, we apply the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy to evaluate the NK-DSGE monetary model summarized in Eq.s (10)-(13), using U.S. quarterly data. Unlike Benati and Surico (2009), we do not force the covariance matrix of the structural disturbances to be diagonal, see, e.g., Kapetanios et al. (2007) Dufour et al. (2013) and Castelnuovo and Fanelli (2013) for similar choices. We fix the discount rate at the value ϱ=0.99 and split the vector of structural parameters θ as $\theta = (\theta_s', \theta_\sigma')'$, where $\theta_s = (\gamma,\ \delta,\ \varkappa,\ \kappa,\ \rho,\ \varphi_\pi,\ \varphi_y,\ \rho_{\tilde{y}},\ \rho_\pi,\ \rho_i)'$ and $\theta_\sigma = vech(\Sigma_{W,u})$. The natural rate of output is approximated with the official measure provided by the Congressional Budget Office (CBO) estimation, following, e.g., Cho and Moreno (2006) and

Castelnuovo and Surico (2010) (see also Table 1 in Gorodnichenko and Ng (2010)). Approximating y_t^p with the CBO time series allows us to treat the 'complete' vector $Z_t = (y_t, \pi_t, i_t, y_t^p)'$ ($n = 4$) as observable. The other variables are the real GDP y_t; the inflation rate π_t, which is the quarterly growth rate of the GDP deflator; and the short-term nominal interest rate i_t, measured by the effective federal funds rate expressed in quarterly terms (averages of monthly values). The data source is the web site of the Federal Reserve Bank of St. Louis.

Our data cover the 'Great Moderation' period, 1985q1-2008q3, hence we have $T = 95$ observations (not including initial lags). The choice of the sample is motivated in our Technical Supplement in detail. We fix the overall nominal level of significance at the 5% level, and the type-I errors of the tests $LR1$, $LR2$ and $LR3$ at the 1%, 2% and 2% levels, respectively. The empirical analysis starts with the estimation of an unrestricted VAR system for Z_t as specified in Eq. (40). We include a constant in the equations (i.e., $d_t = 1$ and μ is a $n \times 1$ constant) because the variables in Z_t are not demeaned prior to estimation. As discussed in Sub-sections 2.1 and 3.1, the system should be driven by a single stochastic trend under the null of the NK-DSGE model, and the cointegration relationships should match the specification of $\beta_0 = \beta_0^b = \zeta$ in Eq. (25). In other words, the variables $\tilde{y}_t = (y_t - y_t^p)$, π_t and i_t should be jointly stationary.

The $LR1$ ($LR1_{\text{seq}}$), $LR2$ and $LR3$ tests are reported in Table 3, Panel 1. We complement the asymptotic p-values of the tests with their bootstrap analogues. Results indicate that the evidence in favour of a single stochastic trend is not clear-cut, but defendable. While the test $LR1$ provides ample support for the hypothesis $r = 3 = p$ ($n - r = 1$) at the 1% level, considering both asymptotic and bootstrap p-values, a different picture emerges from the test $LR1_{\text{seq}}$, which selects $r = \hat{r} = 1$ at the 1% level, irrespective of whether asymptotic or bootstrap p-values are considered. The outcome $r = \hat{r} = 1$ would lead us to conclude that there are three common stochastic trends in the data, two more than expected, and it would be difficult to reconcile such an evidence with a substantial body of work on the drivers of the 'Great Moderation'. Actually, the test $LR1_{\text{seq}}$ has a poor small sample (power) performance, as we have documented in Section 5; hence, it is reasonable to conjecture that the high persistence characterizing the time series π_t and i_t in the period 1985q1-2008q3 induces the test to select two unit roots instead of two stationary roots. Hence, we do not have sufficient evidence to refute the result of the 'one-shot' cointegration test $LR1$. This finding strongly supports the case $r = 3$ ($n - r = 1$). The $LR2$ test provides another piece of evidence in favour of this hypothesis. Indeed, while the asymptotic p-value associated with the $LR2$ test statistic implies rejection, its bootstrap counterpart is equal to 0.04 and does not lead us to reject the structure in Eq. (25) at the 2% level. We therefore consider the last step of the '$LR1 \rightarrow LR2 \rightarrow LR3$' testing strategy.

The last step requires testing the short-run CER. We take the 'partial equilibrium' finite-order VAR representation for $W_t = (\tilde{y}_t,\, \pi_t,\, i_t)'$ in Eq.s (3)-(6) directly to the data. To compute the $LR3$ test, the VAR system (3) is estimated unrestrictedly and under the CER in Eqs. (4)-(6), using the ML algorithm summarized in the Technical Supplement. We split θ_s as $\theta_s = (\theta_g',\, \theta_{ng}')'$, where $\theta_g = (\delta,\, \varkappa,\, \kappa,\, \varphi_\pi,\, \varphi_y)'$ and $\theta_{ng} = (\gamma,\, \rho,\, \rho_{\tilde{y}},\, \rho_\pi,\, \rho_i)'$, and combine a grid-search approach for the elements of θ_g, which are notoriously difficult to estimate through non-Bayesian techniques, with a numerical Newton-type estimation approach for the elements of θ_{ng}. The ML estimate of θ_σ is obtained indirectly, given the estimate of θ_s. Estimation results are summarized in Table 3, Panel 2.

[Table 3 about here.]

The p-value associated with the $LR3$ test is equal to 0.022, while its bootstrap analogue is 0.80; hence, we do not reject the CER implied by the NK-DSGE model at the 2% level. The point estimates of the structural parameters turn out to be quite similar to those found in a variety of contributions in the literature, hence we do not discuss these results in detail. A note of caution is needed for the parameters of the policy reaction function. As we have learned from the Monte Carlo experiment, it is extremely difficult to estimate these parameters precisely in small samples. Weak identification of the parameters might be an important concern; see Mavroeidis (2010), Castelnuovo and Fanelli (2013) and the Technical Supplement. As is well known, weak identification issues may affect the asymptotic (and bootstrap) distribution of the estimators and tests commonly used, hence the evidence stemming from the $LR3$ test should be taken with caution.

Overall, the '$LR1 \to LR2 \to LR3$' testing strategy does not lead to rejection of the NK-DSGE model in Eq.s (10)-(13) at the 5% nominal level, during the 'Great Moderation' period 1985q1-2008q3, with the caveats discussed above.

7 Concluding remarks

DSGE models are interpreted as inherently misspecified systems, hence it is often claimed that testing their implied restrictions is an exercise which is inevitably destined to fail. According to this interpretation, only Bayesian methods are meaningful and can successfully be applied in empirical work. Our paper has shown that a 'frequentist' VAR-based evaluation approach can provide interesting insights with small-scale NK-DSGE models. We have proposed the '$LR1 \to LR2 \to LR3$' testing strategy as a conditional sequence of likelihood-ratio tests which evaluates the long-run and short-run restrictions implied by the NK-DSGE model jointly through a multiple hypothesis testing exercise.

We derive three main lessons from our analysis. First, if the information stemming from the individual tests is used constructively, our approach can be exploited to rectify/modify the structural equations when misspecification is detected. Second, an important cause of rejection may be the poor performance of the tests in small samples. The risk of falsely rejecting the structural model in applied work can be reduced considerably by considering the bootstrap counterpart of the '$LR1 \rightarrow LR2 \rightarrow LR3$' procedure. Third, in samples of the size typically available to practitioners, the weak identification phenomenon of some of the structural parameters is a concrete possibility that deserves attention. However, despite the highly constrained nature of the model, a properly conducted testing approach is not necessarily destined to lead one to reject the model, but will most likely lead to better models and better understanding of the macro economy.

References

An, S. and F. Schorfheide (2007). Bayesian analysis of DSGE models. *Econometric Reviews*, *26*(2-4), 113–172.

Bårdsen, G. and L. Fanelli (2013). Technical supplement to "Frequentist evaluation of small DSGE models". Available at http://www.unibo.it/..

Bekaert, G., S. Cho and A. Moreno (2010). New Keynesian macroeconomics and the term structure. *Journal of Money Credit and Banking*, *42*, 33–61.

Benati, L. and P. Surico (2009). VAR analysis and the great moderation. *American Economic Review*, *99*(4), 1636–1652.

Bianchi, F. (2012). Evolving monetary/fiscal policy mix in the United States. *American Economic Review*, *102*(3), 167–172.

Binder, M. and M. H. Pesaran (1995). Multivariate rational expectations models and macroeconomic modelling: A review and some new results. In Pesaran, M. and M. Wickens (eds.), *Handbook of Applied Econometrics*, 139–187. Blackwell Publishing Ltd, Oxford.

Boswijk, H. P., G. Cavaliere, A. Rahbek and A. M. Taylor (2013). Inference on co-integration parameters in heteroskedastic vector autoregressions. Discussion papers, Department of Economics, University of Copenhagen.

Broze, L., C. Gourieroux and A. Szafarz (1990). *Reduced Forms of Rational Expectations Models*. Harwood fundamentals of pure and applied economics. Harwood Academic Publisher.

Campbell, J. and R. Shiller (1987). Cointegration and tests of present-value models. *Journal of Political Economy*, *95*, 1052–1088.

Canova, F. and F. Ferroni (2012). The dynamics of US inflation: Can monetary policy explain the changes? *Journal of Econometrics*, *167*(1), 47–60.

Canova, F., M. Finn and A. Pagan (1994). Evaluating a real business cycle model. In Hargreaves, C. P. (ed.), *Nonstationary Time Series Analysis and Cointegration*, Advanced Texts in Econometrics, 225–255. Oxford University Press, Oxford.

Carlstrom, C. T., T. S. Fuerst and M. Paustian (2009). Monetary policy shocks, Choleski identification, and DNK models. *Journal of Monetary Economics*, *56*(7), 1014 – 1021.

Castelnuovo, E. and L. Fanelli (2013). Monetary policy indeterminacy and identification failures in the U.S.: Results from a robust test. Working Paper N. 163, Dipartimento di Scienze Economiche ed Aziendali, Marco Fanno.

Castelnuovo, E. and S. Nisticò (2010). Stock market conditions and monetary policy in a DSGE model for the U.S. *Journal of Economic Dynamics and Control*, *34*(9), 1700–1731.

Castelnuovo, E. and P. Surico (2010). Monetary policy, inflation expectations and the price puzzle. *Economic Journal*, *120*, 1262–1283.

Cavaliere, G., A. Rahbek and M. Taylor (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica*, *80*, 1721–1740.

Chang, Y., T. Doh and F. Schorfheide (2007). Non-stationary hours in a DSGE model. *Journal of Money, Credit and Banking*, *39*, 1357–1373.

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

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

Christiano, L. J. (2007). Comment. *Journal of Business & Economic Statistics*, *25*(2), 143–151.

Clarida, R., J. Galí and M. Gertler (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics*, *115*, 147–180.

De Grauwe, P. (2010). The scientific foundation of dynamic stochastic general equilibrium (DSGE) models. *Public Choice*, *144*, 413–443.

DeJong, D. N. and C. Dave (2011). *Structural Macroeconometrics*. Princeton University Press, 2nd edn.

Del Negro, M., F. Schorfheide, F. Smets and R. Wouters (2007). On the fit of new Keynesian models. *Journal of Business & Economic Statistics*, *25*(2), 123–143.

Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys*, *12*(5), 573–593.

Dufour, J.-M., L. Khalaf and M. Kichian (2013). Identification-robust analysis of DSGE and structural macroeconomic models. *Journal of Monetary Economics*, *60*(3), 340–350.

Estrella, A. and J. C. Fuhrer (2002). Dynamic inconsistencies: Counterfactual implications of a class of rational-expectations models. *American Economic Review*, *92*(4), 1013–1028.

Estrella, A. and J. C. Fuhrer (2003). Monetary policy shifts and the stability of monetary policy models. *The Review of Economics and Statistics*, *85*(1), 94–104.

Fanelli, L. (2012). Determinacy, indeterminacy and dynamic misspecification in linear rational expectations models. *Journal of Econometrics*, *170*(1), 153–163.

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

Fuhrer, J. C. and G. D. Rudebusch (2004). Estimating the Euler equation for output. *Journal of Monetary Economics*, *51*(6), 1133–1153.

Fukač, M. and A. Pagan (2010). Limited information estimation and evaluation of DSGE models. *Journal of Applied Econometrics*, *25*(1), 55–70.

Galí, J. (2008). *Monetary Policy, Inflation and the Business Cycle*. Princeton University Press, Princeton.

Gorodnichenko, Y. and S. Ng (2010). Estimation of DSGE models when the data are persistent. *Journal of Monetary Economics*, *57*(3), 325–340.

Guerron-Quintana, P., A. Inoue and L. Kilian (2013). Frequentist inference in weakly identified dynamic stochastic general equilibrium models. *Quantitative Economics*, *4*(2), 197–229.

Hansen, L. P. and T. J. Sargent (1980). Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, *2*, 7–46.

Hansen, L. P. and T. J. Sargent (1981). Linear rational expectations models for dynamically interrelated variables. In Lucas, R. and T. Sargent (eds.), *Rational Expectations and Econometric Practise*, 127–156. University of Minnesota Press, Minneapolis.

Ireland, P. (2004). Technology shocks in the new Keynesian model. *Review of Economics and Statistics*, *86*, 923–936.

Johansen, S. (1996). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model*. Oxford University Press, Oxford, 2nd edn.

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

Johansen, S. and A. R. Swensen (1999). Testing exact rational expectations in cointegrated vector autoregressive models. *Journal of Econometrics*, *93*, 73–91.

Juselius, M. (2011). Testing steady-state restrictions of linear rational expectations models when data are highly persistent. *Oxford Bulletin of Economics and Statistics*, *73*(3), 315–334.

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

King, R. G., C. I. Plosser, J. H. Stock and M. W. Watson (1991). Stochastic trends and economic fluctuations. *The American Economic Review*, *81*(4), 819–840.

Komunjer, I. and S. Ng (2011). Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica*, *79*(6), 1995–2032.

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

Lubik, A. and F. Schorfheide (2004). Testing for indeterminacy: An application to US monetary policy. *American Economic Review*, *94*, 190–217.

Lütkepohl, H. and H. Claessen (1997). Analysis of cointegrated VARMA processes. *Journal of Econometrics*, *80*, 223–239.

Mavroeidis, S. (2010). Monetary policy rules and macroeconomic stability: Some new evidence. *American Economic Review*, *100*, 491–503.

Pesaran, M. H. and R. Smith (2011). Beyond the DSGE straitjacket. *The Manchester School*, *Supplement*, 1–38.

Ravenna, F. and C. E. Walsh (2006). Optimal monetary policy with the cost channel. *Journal of Monetary Economics*, *53*(2), 199–216.

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

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, *15*, 645–670.

Schorfheide, F. (2011). Estimation and evaluation of DSGE models: progress and challenges. *NBER*, *Working Paper 16781*.

Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, *97*(3), 586–606.

Söderlind, P. and A. Vredin (1996). Applied cointegration analysis in the mirror of macroeconomic theory. *Journal of Applied Econometrics*, *11*(4), 363–81.

Stock, J. and M. Watson (1988). Testing for common trends. *Journal of the American Statistical Association*, *83*, 1097–1107.

Uhlig, H. (1999). A toolkit for analyzing nonlinear dynamic stochastic models easily. In Marimon, R. and A. Scott (eds.), *Computational Methods for the Study of Dynamic Economies*. Oxford University Press, Oxford.

Wallis, K. F. (1980). Econometric implications of the rational expectations hypothesis. *Econometrica*, *48*(1), 49–73.

Woodford, M. (2003). *Interest and Prices*. Princeton University Press, Princeton.

Appendix A Size properties of the testing strategy

In this Appendix we discuss the asymptotic size of the 'LR1→LR2 →LR3' testing strategy. Denote with $LR_{i,T}$, $i = 1, 2, 3$, the three LR test statistics of the sequence, and let ψ_i be the nominal significance level (type-I error) pre-fixed for the i-th test; moreover, let $\psi_{i,T} = \mathbb{P}^{H_0}_{i,T}(LR_{i,T} \geq cr^{\psi_i}_{i,T})$ be the exact size of the i-th test based on a sample of size T, where $\mathbb{P}^{H_0}_{i,T}(\cdot)$ is the probability measure associated with the (marginal) null distribution of $LR_{i,T}$ and $cr^{\psi_i}_{i,T}$ is the corresponding critical value at nominal level ψ_i. Under the null H_0 in Eq. (11), the three tests, individually considered, are correctly sized; in particular, they satisfy the condition $\psi_{i,\infty}$ =lim sup $\psi_{i,T} = \psi_i$, $i = 1, 2, 3$, where $\psi_{i,\infty}$ is the asymptotic size of the i-th test and 'lim sup' is intended for $T \to \infty$. Let $\mathbb{P}^{H_0}_{1,2,T}(\cdot\ ;\ \cdot)$ and $\mathbb{P}^{H_0}_{2,3,T}(\cdot\ ;\ \cdot)$ be the (joint) probability measures associated with the null distributions of the test statistics $LR_{1,T}$ and $LR_{2,T}$ and the test statistics $LR_{2,T}$ and $LR_{3,T}$, respectively. It turns out that the overall asymptotic size of the testing strategy is given by ψ_∞ =lim sup ψ_T, where

$$\psi_T = \mathbb{P}^{H_0}_{1,T}(LR_{1,T} \geq cr^{\psi_1}_{1,T}) + \mathbb{P}^{H_0}_{1,2,T}(LR_{1,T} < cr^{\psi_1}_{1,T}\ ;\ LR_{2,T} \geq cr^{\psi_2}_{2,T}\ ) + \mathbb{P}^{H_0}_{2,3,T}(LR_{2,T} < cr^{\psi_2}_{2,T}\ ;\ LR_{3,T} \geq cr^{\psi_3}_{3,T}). \tag{A42}$$

The first addend of Eq. (A42) captures the probability that the test LR1 incorrectly rejects the cointegration rank in a sample of size T ; the second addend captures the joint probability that the LR2 test incorrectly rejects the structure of the cointegration matrix; the LR1 test correctly selects the cointegration rank in a sample of size T; and, finally, the last addend captures the joint probability that the LR3 test incorrectly rejects the CER and the LR2 correctly rejects the structure of the cointegration matrix in a sample of size T. By using the inequalities $\mathbb{P}^{H_0}_{1,2,T}(LR_{1,T} < cr^{\psi 1}_{1,T}; LR_{2,T} \geq cr^{\psi 2}_{2,T}) \leq \psi_{2,T}$ and $\mathbb{P}^{H_0}_{2,3,T}(LR_{2,T} < cr^{\psi_2}_{2,T}\ ;\ LR_{3,T} \geq cr^{\psi_3}_{3,T}) \leq \psi_{3,T}$, the asymptotic size ψ_∞ is such that

$$\psi_\infty \leq \psi_{1,\infty} + \psi_{2,\infty} + \psi_{3,\infty} = \psi_1 + \psi_2 + \psi_3. \tag{A43}$$

This result suggests that, in empirical analyses, it is convenient to fix the overall nominal significance level of the procedure at the level ψ =($\psi_1 + \psi_2 + \psi_3$).

Table 1: Monte Carlo results: size of the testing strategy and ML estimates of the structural parameters.

Panel 1: Empirical size of the '$LR1 \to LR2 \to LR3$' testing strategy					
Tests	$T = 100$		$T = 200$		$T = 500$
	Asymptotic	Bootstrap	Asymptotic	Bootstrap	Asymptotic
$LR1_{\psi_1=0.01}$ $(r = 3)$	0.006	0.006	0.009	0.009	0.012
$LR2_{\psi_2=0.02}$ $(\beta_0 = \zeta \mid LR1)$	0.072	0.022	0.043	0.022	0.022
$LR3_{\psi_3=0.02}$ (CER $\mid LR2$)	0.028	0.008	0.019	0.019	0.016
Overall rejection: $\hat{\psi} = \sum_{i=1}^{3} \hat{\psi}_i$	0.106	0.036	0.071	0.050	0.050

Panel 2: ML estimates and s.e. of structural parameters: $\underset{(s.e._{MC}(\hat{\theta}_i))}{\hat{E}_{MC}(\hat{\theta}_i)}$			
True parameters	$T = 100$	$T = 200$	$T = 500$
$\kappa = 0.044$	0.095 (0.111)	0.067 (0.058)	0.049 (0.029)
$\delta = 0.124$	0.149 (0.082)	0.136 (0.053)	0.129 (0.032)
$\gamma = 0.744$	0.740 (0.079)	0.745 (0.053)	0.744 (0.033)
$\omega_f = 0.935$	0.955 (0.180)	0.932 (0.148)	0.912 (0.122)
$\rho = 0.834$	0.826 (0.085)	0.832 (0.060)	0.832 (0.037)
$\varphi_y = 1.146$	1.440 (1.197)	1.356 (0.841)	1.207 (0.390)
$\varphi_\pi = 1.749$	2.436 (1.680)	2.155 (1.206)	1.859 (0.592)
$\rho_y = 0.796$	0.768 (0.141)	0.784 (0.079)	0.789 (0.043)
$\rho_\pi = 0.418$	0.404 (0.205)	0.380 (0.079)	0.362 (0.157)
$\rho_i = 0.404$	0.394 (0.135)	0.402 (0.098)	0.404 (0.062)
$\sigma^2_{\tilde{y}} = 0.055$	0.072 (0.041)	0.062 (0.024)	0.058 (0.013)
$\sigma^2_\pi = 0.391$	0.450 (0.181)	0.429 (0.108)	0.421 (0.079)
$\sigma^2_i = 0.492$	0.515 (0.164)	0.508 (0.106)	0.496 (0.053)
$\sigma^2_{y^p} = 0.020$	0.020 (0.003)	0.020 (0.002)	0.020 (0.001)

NOTES: Results are obtained using $M = 5000$ Monte Carlo replications generated under the null of the NK-DSGE model in Eq.s (10)-(14). Given the initial conditions, the observations $Y_1, ..., Y_T$ are generated from the VAR system (21)-(24) and then transformed into $Z_1, ..., Z_T$ using the restriction $\beta_0 = \beta_0^b = \zeta$ from Eq. (25) and the mapping $Z_t = G\left(\beta_0^b, \tau, 1-L\right)^{-1} Y_t$. For each replication, a sample of $T + 200$ observations is generated and the first 200 observations are then discarded. PANEL 1: empirical rejection frequencies (erf) of the tests $LR1$, $LR2$ and $LR3$ and of the overall '$LR1 \to LR2 \to LR3$' testing strategy; the column 'Asymptotic' reports the erf computed using the asymptotic critical values taken from Doornik (1998); the column 'Bootstrap' reports the erf computed using the bootstrap p-values associated with the tests; the 'one-shot' cointegration rank test $LR1$ evaluates the null of a single stochastic trend versus the alternative of stationary VAR and is computed from a VAR system for Z_t as in Eq. (40) with $\ell = 2$ and no deterministic components; the (iid) bootstrap counterpart of the test $LR1$ is computed using the method discussed in Cavaliere et al. (2012) with $B = 399$ replications; the over-identified cointegrating restrictions test $LR2$ is computed from the error-correction system as in Eq. (21) with $\ell = 2$ and no deterministic components and evaluates whether β_0 has the structure in Eq. (25) and has $12 - 9 = 3$ degrees of freedom; the (iid) bootstrap counterpart of the test $LR2$ is computed using the method discussed in Boswijk et al. (2013) with $B = 399$ replications; the test $LR3$ is computed by estimating a VAR system for Y_t as in Eq. (21) unrestrictedly and under the CER in Eq.s (22)-(24) by the ML algorithm summarized in the Technical Supplement, and has $42 - 14 = 28$ degrees of freedom; the bootstrap p-value for the test $LR3$ is computed with $B = 99$ replications and using the non-parametric analogue of the procedure discussed in Fanelli and Palomba (2011, Section 3), case $t = T$. PANEL 2: Averages of the ML estimates of the structural parameters and Monte Carlo standard errors in parentheses; averages are computed considering only DGPs for which the '$LR1 \to LR2 \to LR3$' testing strategy does not lead to rejection; ML estimates are obtained by maximizing the Gaussian log-likelihood of the VAR system for Y_t, see system Eq. (21), unrestrictedly and under the CER in Eq.s (22)-(24); see Technical Supplement.

Table 2: Simulated and bootstrapped marginal acceptance frequencies of $LR1_{seq}$ and rejection frequencies of the '$LR1_{\text{seq}} \to LR2 \to LR3$' testing strategy of the NK-DSGE model.

Panel 1: Empirical acceptance frequencies of the $LR1_{\text{seq}}$ test					
Tests		$T = 100$		$T = 200$	
		Asymptotic	Bootstrap	Asymptotic	Bootstrap
$LR1_{\text{seq}}$	$r = 0$	0.010	0.036	0.000	0.000
	$r = 1$	0.353	0.445	0.000	0.001
	$r = 2$	0.539	0.437	0.276	0.326
	$r = 3$	0.094	0.078	0.712	0.663
	$r = 4$	0.004	0.004	0.012	0.010
Panel 2: Empirical size of the '$LR1_{\text{seq}} \to LR2 \to LR3$' testing strategy					
$LR1_{\text{seq}\,\psi_1=0.01}$ $(r = 3)$		0.004	0.004	0.012	0.010
$LR2_{\psi_2=0.02}$ $(\beta_0 = \zeta \mid LR1_{\text{seq}})$		0.140	0.059	0.044	0.022
$LR3_{\psi_3=0.02}$ (CER $\mid LR2$)		0.049	0.014	0.022	0.011
Overall rejection: $\hat{\psi} = \sum_{i=1}^{3} \hat{\psi}_i$		0.193	0.077	0.078	0.043

NOTES: Results are obtained using $M = 5000$ Monte Carlo replications generated as detailed in the notes of Table 1. Panel 1: The column 'Asymptotic' reports the empirical acceptance frequencies (*eaf*) computed using the asymptotic critical values; the column 'Bootstrap' reports the *eaf* computed using the bootstrap p-values associated with the tests; the (iid) bootstrap counterpart of the test LR1_{seq} is computed following Cavaliere et al. (2012) using $B = 399$ replications. Panel 2: Empirical rejection frequencies of the tests $LR1_{\text{seq}}$, $LR2$ and $LR3$ and of the overall '$LR1_{\text{seq}} \to LR2 \to LR3$' testing strategy. The tests $LR2$ and $LR3$, including their bootstrap counterparts, are computed as detailed in the notes to Table 1.

Table 3: The tests $LR1_{seq}$, LR1, LR2 and LR3, '$LR1_{\text{seq}} \to LR2 \to LR3$' testing strategy and ML estimates of the structural parameters of the NK-DSGE system on U.S. quarterly data, 1985q1-2008q3.

Panel 1: tests of NK-DSGE model				
Tests		Trace	Asymptotic	Bootstrap
$LR1_{\text{seq}}$:	$r = 0$	107.10	0.000	0.000
	$r = 1$	32.33	0.024	0.071
	$r = 2$	15.07	0.056	0.248
$LR1_{\psi_1=0.01}$ $(r = 3)$	$r = 3$	2.43	0.119	0.491
$LR2_{\psi_2=0.02}$ $(\beta_0 = \zeta \mid LR1)$		11.665	0.009	0.040
$LR3_{\psi_3=0.02}$ $(CER \mid LR2)$		17.94	0.022	0.80

Panel 2: ML estimates of structural parameters		
Parameters θ_s:	**Interpretation**	**ML**
γ	AD, forward look. term	0.777 (0.025)
δ	AD, inverse elasticity of sub.	0.030 (0.006)
$\varkappa$	NKPC, indexation	0.014 (0.015)
implied value of $\omega_f = \frac{0.99}{1+0.99\varkappa}$	NKPC, forward-looking	0.977 (0.034)
κ	NKPC, slope	0.083 (0.022)
ρ	Policy rule, smoothing term	0.573 (0.358)
$\varphi_{\tilde{y}}$	Policy rule, react. to out. gap	0.073 (1.145)
φ_{π}	Policy rule, react. to inflation	5.37 (2.47)
$\rho_{\tilde{y}}$	AD, disturbance persist.	0.935 (0.010)
ρ_{π}	NKPC, disturbance persist.	0.875 (0.011)
ρ_i	Policy rule, disturbance persist.	0.810 (0.451)

$$\textbf{Parameters} \quad \theta_\sigma : \qquad \hat{\Sigma}_{W,u} = \begin{pmatrix} \underset{(0.0002)}{0.0145} & \underset{(0.0003)}{-0.0019} & \underset{(0.0018)}{-0.0008} \\ & \underset{(0.0007)}{0.0051} & \underset{(0.0041)}{-0.0223} \\ & & \underset{(0.032)}{0.222} \end{pmatrix}$$

NOTES: Results are obtained from a VAR system for $Z_t = (y_t,\ \pi_t,\ i_t,\ y_t^p)'$ as specified in Eq. (40) with $\ell = 2$, $d_t = 1$ and μ unrestricted. PANEL 1: The column 'Trace' reports the LR cointegration rank Trace statistic; the column 'Asymptotic' reports the p-values of the test computed with asymptotic critical values from Doornik (1998); the column 'Bootstrap' reports the p-values of the test computed with the bootstrap; the 'one-shot' cointegration rank test $LR1$ evaluates the null of a single stochastic trend versus the alternative of a stationary VAR and is highlighted in the fourth row; the bootstrap p-values for the tests $LR1$ and $LR1_{\text{seq}}$ are computed using the method discussed in Cavaliere et al. (2012) with $B = 399$ replications; the test $LR2$ evaluates the over-identification cointegration restrictions in Eq. (25) and has 3 degrees of freedom; the (iid) bootstrap counterpart of the test $LR2$ is computed using the method discussed in Boswijk et al. (2013), with $B = 399$ replications; the test $LR3$ evaluates the CER implied by the NK-DSGE model and has 8 degrees of freedom; the bootstrap p-value for the test $LR3$ is computed with $B = 99$ replications and using the non-parametric analogue of the procedure discussed in Fanelli and Palomba (2011, Section 3), case $t = T$. PANEL 2: ML estimates have been obtained from the finite-order VAR for $W_t^o = (y_t - y_t^p,\ \pi_t,\ i_t)'$ in Eq. (3) by maximizing the Gaussian log-likelihood under the CER in Eq.s (4)-(6) by combining the BFGS method for γ, ρ, $\rho_{\tilde{y}}$, ρ_π and ρ_i with a grid search for δ (range [0.01, 0.20]), $\varkappa$ (range [0.01, 0.10]), κ (range [0.01, 0.10]), $\varphi_{\tilde{y}}$ (range [0.05, 1.50]) and φ_π (range [0.5, 5.50]) (see Technical Supplement); the covariance matrix $\Sigma_{W,u}$ is not diagonal and its elements are estimated indirectly (see Technical Supplement); the variables in W_t^o have been preliminarily demeaned; ML estimates are robust to different choices of the initial values used for γ, ρ, $\rho_{\tilde{y}}$, ρ_π and ρ_i; asymptotic standard errors are reported in parentheses below estimates; 'AD' stands for aggregate demand; 'NKPC' stands for New Keynesian Phillips curve.