

# BOOTSTRAPPING DSGE MODELS

BY GIOVANNI ANGELINI<sup>†</sup>, GIUSEPPE CAVALIERE<sup>†</sup> AND LUCA FANELLI<sup>†\*</sup>

August 29, 2017

## ABSTRACT

This paper explores the potential of bootstrap methods in the empirical evaluation of dynamic stochastic general equilibrium (DSGE) models and, more generally, in linear rational expectations models featuring unobservable (latent) components. We consider two dimensions. First, we provide mild regularity conditions that suffice for the bootstrap Quasi-Maximum Likelihood (QML) estimator of the structural parameters to mimic the asymptotic distribution of the QML estimator. Consistency of the bootstrap allows to keep the probability of false rejections of the cross-equation restrictions under control. Second, we show that the realizations of the bootstrap estimator of the structural parameters can be constructively used to build novel, computationally straightforward tests for model misspecification, including the case of weak identification. In particular, we show that under strong identification and bootstrap consistency, a test statistic based on a set of realizations of the bootstrap QML estimator approximates the Gaussian distribution. Instead, when the regularity conditions for inference do not hold as e.g. it happens when (part of) the structural parameters are weakly identified, the above result is no longer valid. Therefore, we can evaluate how close or distant is the estimated model from the case of strong identification. Our Monte Carlo experimentations suggest that the bootstrap plays an important role along both dimensions and represents a promising evaluation tool of the cross-equation restrictions and, under certain conditions, of the strength of identification. An empirical illustration based on a small-scale DSGE model estimated on U.S. quarterly observations shows the practical usefulness of our approach.

KEYWORDS: Bootstrap, Cross-equation restrictions, DSGE, QLR test, State space model, Weak identification.

JEL CLASSIFICATION: C32, C51, E30, E50.

---

<sup>†</sup>School of Economics, Management and Statistics, University of Bologna, Italy. \*Corresponding author. We thank Anders Rahbek, Peter Boswijk and Luca Sala for helpful comments. Funding from MIUR Grant PRIN-2010/2011, prot. 2010RHAHPL 003 and RFO grants from the University of Bologna is gratefully acknowledged. Correspondence to: Luca Fanelli, School of Economics, Management and Statistics, University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy; email: luca.fanelli@unibo.it.

# 1 INTRODUCTION

Dynamic stochastic general equilibrium (DSGE) models are linear(ized) rational expectations models currently used by central banks and academicians to evaluate macroeconomic policies and to predict the stance of the business cycle. These models are stylized representations of the economy and are misspecified in several dimensions (An and Schorfheide, 2007). They are typically treated by econometricians as restricted but parametrically incomplete representations of the actual data (Diebold *et al.* 1998). DSGE models imply highly nonlinear restrictions on the state space representation they generate. These restrictions, denoted ‘cross-equation restrictions’ (CER) hereafter, are the ‘hallmark’ of rational expectations models (Hansen and Sargent, 1980, 1981) and represent the ‘natural’ metric through which these models should be evaluated empirically (Hansen, 2014). In the frequentist setting, DSGE models are usually rejected when evaluated through the CER.

Following the original intuition of Mankiw and Shapiro (1987) who focused on orthogonality restrictions, Bekaert and Hodrick (2001) have shown that in rational expectations models, commonly employed tests of the CER based on standard asymptotic approximations may lead to severe size distortions and power losses in finite samples. In particular, the empirical size of these tests tends to exceed markedly the chosen significance level, inducing practitioners to falsely conclude that their models are too simple to capture the complex probabilistic nature of the data.

As is well known, the bootstrap, when correctly implemented, is an important device for improving upon the finite sample size properties of asymptotic tests (Horowitz, 2001).<sup>1</sup> Cho and Moreno (2006) and Bårdsen and Fanelli (2015) have exploited bootstrap methods for testing the CER implied by small-scale New-Keynesian DSGE models, documenting substantial finite-sample size improvements under the null. Unfortunately, these authors consider the restrictive case where all endogenous variables of the system are observed (or can be easily proxied by observables) and their dynamic properties can be approximated by finite-order vector autoregressions (VAR), while it is now well known that VAR representations with a finite number of lags are the exception rather than the rule in DSGE models (Ravenna, 2007; Franchi and Vidotto, 2013; Franchi and Paruolo 2015). More crucially, both Cho and Moreno (2006) and Bårdsen and Fanelli (2015) do not assess whether the regularity conditions which permit standard inference – generically denoted with the term ‘strong identification’ in the following – are valid. Conversely, the recent literature suggests that ‘sample’ or ‘weak’ identification problems, defined as cases where the likelihood function of the system does not fully respect the usual regularity conditions necessary for standard inference, are an important issue in DSGE models. See Canova and Sala (2009), Dufour *et al.* (2009, 2013), Kleibergen and Mavroeidis (2009), Mavroeidis (2005, 2010), Guerron-Quintana *et al.* (2013), An-

---

<sup>1</sup>In macroeconomic analysis, bootstrap methods are typically used either to build confidence bands for impulse response functions computed in structural vector autoregressions (Kilian, 1998), or to obtain confidence intervals for the structural parameters of dynamic macro models (Cho and Moreno, 2006). They can also be conveniently use to draw inferences on the long run relationships among economic variables, see Cavaliere *et al.* (2012), Cavaliere *et al.* (2015a) and Boswijk *et al.* (2015). Only seldom they have been used to improve the small sample performance of tests of the CER in linear(ized) rational expectations models.

Andrews and Mikusheva (2015), Qu (2014) and Castelnuovo and Fanelli (2015), among others.

In the current literature, little is known about the asymptotic properties of bootstrap methods in DSGE models and how to implement these methods to test the CER. Moreover, little is known on how the bootstrap performs in practice and nothing is known about the performance of the bootstrap in weakly identified DSGE models. In this paper we fill the gap and explore the potential of bootstrap methods in the empirical evaluation of DSGE models, along at least two dimensions.

First, we develop a (time-domain) bootstrap-based approach for quasi-likelihood ratio (QLR) tests of the CER implied by DSGE models and, more generally, by linear(ized) rational expectations models involving unobservable (latent) components. To do so, we extend Stoffer and Wall's (1991) nonparametric bootstrap approach for state space models to the case of DSGE models. In the case of strong identification, we generalize and strengthen Stoffer and Wall's (1991) results. In particular, we relax the conditions used by Stoffer and Wall's (1991) to show that the bootstrap quasi-maximum likelihood (QML) estimator of the structural parameters replicates the asymptotic distribution of the QML estimator, and prove formally that the restricted bootstrap (i.e., with the null hypothesis under investigation being imposed in estimation) is consistent. In this scenario, the asymptotic distribution of the QML estimator of the structural parameters can be estimated accurately by the bootstrap. Bootstrap confidence intervals for the structural parameters and for impulse responses can be built and interpreted in the conventional way. Importantly, not only the (either standard or bootstrap) QLR test for the CER is asymptotically pivotal and chi-square distributed, but the bootstrap tends to reduce the discrepancy between actual and nominal probabilities of type-I error. It turns out that the bootstrap in DSGE models (and, generally, in state space models) has the potential to attenuate the over-rejection phenomenon that characterizes tests of the CER when first-order asymptotic approximations are used.

Second, inspired by the work of Zhan (2014) who uses the bootstrap in instrumental variable regressions to detect weak instruments, we show the novel result that the bootstrap can also be implemented in order to assess how far the estimated DSGE model is from the case of strong identification. Our starting point is that, when the regularity conditions for standard likelihood-based inference are not met – as it happens e.g. under weak identification in some part of the parameter space – the asymptotic distribution of the QML estimator of the structural parameters is no longer Gaussian. Andrews and Cheng (2012) develop a general non-standard asymptotic theory for nonlinear models and extremum estimators which covers cases where lack of identification and/or weak identification occurs in part of the parameter space. Unfortunately, they assume the validity of a parameterization of the model which is not always easy to check for all DSGE models of interest. Despite this, we show that even in these cases the bootstrap is still useful and can potentially be used to detect some types of misspecification, including weak identification.<sup>2</sup> A simple, ‘descriptive’ indicator of weak identification is

---

<sup>2</sup>Because in weakly identified DSGE model we do not have asymptotically pivotal estimators and test statistics, the common wisdom is that the bootstrap ‘does not work’. While the fact that the bootstrap does not work under non-standard conditions is not always necessarily true (see e.g. Cavaliere *et al.* 2015b for an example), our intuition is that the distribution of the bootstrap QML estimator of the

given by empirical ‘distance’ between the analytical standard errors associated with the QML estimates of the structural parameters and the bootstrap standard errors. Indeed, in strongly identified models, the bootstrap standard errors estimate consistently, conditionally on the original data, the asymptotic standard errors, while this result is no longer true when standard inference does not apply. More important, we propose a test statistic, based on an arbitrary number, say  $\bar{B}$ , of realizations of the bootstrap QML estimator, to test the hypothesis that the model is strongly identified against a generic alternative of failure of the regularity conditions for standard inference. The test statistic is essentially a standard normality test statistic, which is therefore straightforward to compute in practice. We derive sufficient conditions on the number of bootstrap repetitions  $\bar{B}$  for the test statistic to have a well-defined asymptotic distribution under the null of strong identification. In summary, we can use the bootstrap replicates of the QML estimator of the structural parameters together with simple normality tests in order to evaluate how close or distant is the estimated DSGE model from the case of strong identification.

In practice, our approach requires (i) the estimation of the state space representation associated with the DSGE model on the original sample, with and without the CER, (ii) the computation of the QLR test for the CER, and (iii) the application of a (non)parametric bootstrap algorithm which imposes the null of the CER. Our bootstrap algorithm provides: (a) the bootstrap p-value associated with the QLR test and bootstrap standard errors for the estimated structural parameters, (b) an estimate of the distribution of the bootstrap QML estimator of the structural parameters and (c) an estimate of the distribution of the bootstrap QML estimator of the parameters of the state space representation of the DSGE model without the CER imposed. The bootstrap distributions in (b) and possibly in (c) are then analyzed by standard normality tests. We discuss how the information in (a)-(b)-(c) can be processed to analyze the DSGE model. In particular, we show that the proper combination of the analysis in (a) and the diagnostics in (b) and (c) provides an exhaustive evaluation of the features of the estimated model.

We investigate the usefulness of our approach by a set of Monte Carlo experiments using An and Schorfheide’s (2007) small-scale monetary DSGE model as laboratory. Further Monte Carlo studies based on the workhorse autoregressive moving average process of order one (ARMA(1,1)) and Andrews and Mikusheva’s (2015) DSGE model are confined in a Technical Supplement (Angelini *et al.*, 2016). We also provide an empirical illustration based on the estimation of An and Schorfheide’s (2007) model on U.S. quarterly data. Overall, our results lead us to three considerations. First, when all regularity conditions for standard inference are at work, the bootstrap improves the finite sample properties of estimators and tests in DSGE models and contributes to reduce the risk of type I errors. Second, a practitioner who combines our ‘descriptive’ indicator of weak identification (based on the distance between asymptotic and bootstrap standard errors) with our novel bootstrap-based misspecification test can detect weak identification reasonably well. Third, when the failure of the conditions for standard asymptotic inference coincide with the weak identification of a subset of the structural parameters is informative and useful also under weak identification.

structural parameters, we find that the QLR test for the CER can be undersized or oversized depending on the specific model at hand, albeit not dramatically. Remarkably, in these cases the bootstrap still improves upon the asymptotic QLR test, in the sense that the empirical size of the bootstrap QLR test tends to be close the chosen nominal level. Overall, our analysis contrasts the common wisdom that the bootstrap distributions of DSGE estimators or test statistics are not informative in the context of weak identification.

This paper is organized as follows. Section 2 introduces the reference structural DSGE model and the underlying assumptions, and discusses its state space representations. Section 3 focuses on the QLR test for the CER. Section 3.1 introduces the testing problem and Section 3.2 characterizes the concepts of strong and weak identification we refer in this paper. Section 4 summarizes our bootstrap approach to DSGE models. Section 4.1 presents the bootstrap algorithm, Section 4.2 proves its asymptotic validity and Section 4.3 derives our novel bootstrap-based test to detect strong/weak identification. Section 4.4 frames our approach within the existing literature. Section 5 explores the finite sample performance of our approach by some Monte Carlo simulations. Section 6 illustrates how our approach works on actual data by taking An and Schorfheide's (2007) DSGE monetary model to U.S. quarterly data. Section 7 contains some concluding remarks. Appendix A contains technical details and proofs. The associated Technical Supplement (Angelini *et al.*, 2016) complements the results of the paper in several dimensions.

## 2 STRUCTURAL MODEL, STATE SPACE REPRESENTATIONS AND ASSUMPTIONS

Let  $Z_t := (Z_{1,t}, Z_{2,t}, \dots, Z_{n_z,t})'$  be a  $n_z \times 1$  vector of endogenous, possibly unobserved variables at time  $t$ , which can be interpreted as deviation from corresponding steady state values. We assume that, after log-linearization and, for  $t = 1, \dots, T$ , the structural form of the DSGE model reads as

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

$$\eta_t = R \eta_{t-1} + \omega_t \quad , \quad \omega_t \sim \text{WN}(0, \Sigma_\omega). \quad (2)$$

In eq.s (1)-(2),  $\Gamma_i := \Gamma_i(\theta)$ ,  $i = 0, f, b$  are  $n_z \times n_z$  matrices whose elements depend on the vector of structural parameters  $\theta$ ,  $\Pi := \Pi(\theta)$  is an  $n_z \times n_\omega$  matrix of full-column rank ( $n_\omega \leq n_z$ ), whose elements may depend on  $\theta$  but which essentially selects the shocks that enter the equations,  $\eta_t$  is a  $n_\omega \times 1$  vector of autoregressive disturbances,  $R := R(\theta)$  is an  $n_\omega \times n_\omega$  stable diagonal matrix,  $\omega_t$  is the  $n_\omega \times 1$  vector of white noise structural shocks with covariance matrix  $\Sigma_\omega := \Sigma_\omega(\theta)$ . Here,  $\omega_t$  is adapted to the sigma-field  $\mathcal{F}_t$ , where  $\mathcal{F}_t$  is the agents' information set at time  $t$  and  $E_t Z_{t+1} := E(Z_{t+1} | \mathcal{F}_t)$ . The term  $\omega_t$  will be referred to as the vector of fundamental structural shocks, and its covariance matrix  $\Sigma_\omega$  can be either diagonal or non-diagonal. The initial condition  $Z_0$  are treated as given. Finally, the true value of  $\theta$  is denoted by  $\theta_0$  and is assumed to be an interior point of the compact parameter space  $\mathcal{P}$ .

The multivariate linear(ized) rational expectations model described by system (1)-(2) nests a large class of DSGE models currently used in policy and business cycle analysis. Our starting point is system (1)-(2) rather than its state space representation (which is derived below), essentially because our approach can cover a general class of linear rational expectations models used in macroeconomics and finance. The next example considers the prototype monetary DSGE model discussed in the literature, see Komunjer and Ng (2011), Qu and Tkachenko (2012) and Qu (2014), among others, which will be analyzed in detail in the next sections.

EXAMPLE 1 (AN AND SCHORFHEIDE'S (2007) MODEL) *With a slight change of notation, An and Schorfheide's (2007) DSGE model is given by the equations:*

$$x_t = E_t x_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau}(r_t - E_t \pi_{t+1} - E_t z_{t+1}), \quad (3)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(x_t - g_t), \quad (4)$$

$$c_t = x_t - g_t, \quad (5)$$

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)\psi_1 \pi_t + (1 - \rho_r)\psi_2(x_t - g_t) + \varepsilon_{r,t}, \quad \varepsilon_{r,t} \sim WN(0, \sigma_r^2) \quad (6)$$

$$g_t = \rho_g g_{t-1} + \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim WN(0, \sigma_g^2) \quad (7)$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t}, \quad \varepsilon_{z,t} \sim WN(0, \sigma_z^2). \quad (8)$$

eq. (3) is a forward-looking output-gap equation and  $x_t$  is the output gap; eq. (4) is a purely forward-looking New-Keynesian Phillips Curve (NKPC) with slope  $\kappa := \frac{\tau(1-\nu)}{\nu\phi^*\bar{\pi}^2}$  and  $\pi_t$  is the inflation rate; eq. (5) is a consumption equation and  $c_t$  is consumption; eq. (6) is the monetary policy rule and  $r_t$  is the policy rate; finally, eq.s (7)-(8) maintain that the aggregate supply ( $g_t$ ) and demand ( $z_t$ ) disturbances are autoregressive processes. The vector of structural parameters is given by  $\theta := (\tau, \beta, \kappa, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_z^2, \sigma_g^2, \sigma_r^2)'$ ,  $\dim(\theta) = 11$ . We refer to An and Schorfheide (2007) for a derivation and discussion of the system in eq.s (3)-(8). It is seen that the consumption equation (5) does not bear any independent information on  $\theta$  other than contained in the other equations, hence it can be dropped from the structural equations without any loss of information. In terms of the notation in eq.s (1)-(2) we have:  $Z_t := (x_t, \pi_t, r_t)'$ , ( $n_z = 3$ ),  $\eta_t := (z_t, g_t, \varepsilon_{r,t})'$ ,  $\omega_t := (\varepsilon_{z,t}, \varepsilon_{g,t}, \varepsilon_{r,t})'$  ( $n_\omega = 3$ ) and

$$\begin{aligned} \Gamma_0 &:= \begin{pmatrix} 1 & 0 & \tau^{-1} \\ -\kappa & 1 & 0 \\ -(1-\rho_r)\psi_2 & -(1-\rho_r)\psi_1 & 1 \end{pmatrix}, \quad \Gamma_f := \begin{pmatrix} 1 & \tau^{-1} & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \Gamma_b &:= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho_r \end{pmatrix}, \quad \Pi := \begin{pmatrix} \tau^{-1}\rho_e & (1-\rho_g) & 0 \\ 0 & -\kappa & 0 \\ 0 & -(1-\rho_r)\psi_2 & 1 \end{pmatrix} \\ R &:= \begin{pmatrix} \rho_e & 0 & 0 \\ 0 & \rho_g & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Sigma_\omega := \begin{pmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_g^2 & 0 \\ 0 & 0 & \sigma_r^2 \end{pmatrix}. \end{aligned} \quad (9)$$

The derivation of the rational expectations solution of system (1)-(2) and the associated state space representation in ‘minimal form’ is discussed in detail in the Technical Supplement. We assume that a unique and stable solution exists.

ASSUMPTION 1 (DETERMINACY) *Given system (1)-(2),  $\theta_0$  is an interior point of the determinacy region  $\mathcal{P}^D$  of the compact parameter space  $\mathcal{P}$ ,  $\mathcal{P}^D \subseteq \mathcal{P}$ .*

Assumption 1 implies that in correspondence of  $\theta_0$ , the stochastic process that solves system (1)-(2) is covariance stationary. We strengthen this condition by postulating that it is also ergodic. Moreover, it depends neither on extra parameters other than  $\theta$ , nor on additional stochastic shocks other than the fundamental shocks  $\omega_t$ , see Lubik and Schorfheide (2003, 2004), Fanelli (2012) and Castelnovo and Fanelli (2015). We prefer to rule out indeterminate equilibria because the occurrence of extra parameters and shocks other than fundamental shocks complicates identification issues considerably.<sup>3</sup> Let  $y_t := (y_{1,t}, y_{2,t}, \dots, y_{n_y,t})'$  be the  $n_y \times 1$  vector of (de-measured) observed variables.

Under Assumption 1, the so-called ABCD form associated with the determinate solution of the DSGE model in eq.s (1)-(2) can be expressed in the form (see Technical Supplement)

$$\begin{matrix} Z_{m,t} \\ n_m \times 1 \end{matrix} = \begin{matrix} A(\phi_\theta) \\ n_m \times n_m \end{matrix} \begin{matrix} Z_{m,t-1} \\ n_m \times 1 \end{matrix} + \begin{matrix} B(\phi_\theta) \\ n_m \times n_\omega \end{matrix} \begin{matrix} \omega_t \\ n_\omega \times 1 \end{matrix} \quad (10)$$

$$\begin{matrix} y_t \\ n_y \times 1 \end{matrix} = \begin{matrix} C(\phi_\theta) \\ n_y \times n_m \end{matrix} \begin{matrix} Z_{m,t-1} \\ n_m \times 1 \end{matrix} + \begin{matrix} D(\phi_\theta) \\ n_y \times n_\omega \end{matrix} \begin{matrix} \omega_t \\ n_\omega \times 1 \end{matrix} \quad (11)$$

where  $Z_{m,t}$  is the  $n_m$ -dimension sub-vector of  $Z_t$  that contains the candidate ‘minimal’ states of the system,  $A(\phi_\theta)$ ,  $B(\phi_\theta)$ ,  $C(\phi_\theta)$  and  $D(\phi_\theta)$  are matrices of parameters that depend on  $\theta$  through the relationship  $\phi_\theta = g(\theta)$ , where  $g(\cdot)$  is a nonlinear differentiable vector function and  $\phi_\theta$  can be interpreted as a vector of ‘reduced form’ coefficients that depend nonlinearly on  $\theta$  under the CER. Obviously,  $\phi_{\theta_0} = g(\theta_0)$ . We come back on the role of  $\phi_\theta$  at the end of this section and in Section 3. Assumption 1 ensures that the matrix  $A(\phi_\theta) = A(g(\theta))$  in eq. (10) is stable<sup>4</sup>, but this does not necessarily imply that  $A(\phi_\theta)$  is also invertible. In most applications, however, it is found that  $A(\phi_\theta)$  is invertible. Throughout the paper we also impose, without loss of generality, that  $A(\phi_\theta)$  is non-singular.

The ABCD form in eq.s (10)-(11) is the candidate minimal state space representation associated with the DSGE model. It is minimal if and only if the system is controllable and observable, i.e. if the following rank conditions are valid

$$\text{rank}(C^0(\phi_{\theta_0})) = n_m = \text{rank}(O^0(\phi_{\theta_0})) \quad (12)$$

<sup>3</sup>In principle, however, our approach can be extended to the case of DSGE models with indeterminate equilibria. All that we need is the associated (minimal) identified state space representation of the system, see e.g. Lubik and Schorfheide (2004).

<sup>4</sup>Throughout the paper we use the term ‘stable’ to denote a matrix that has all eigenvalues inside the unit circle in the complex plane.

where

$$C^0(\phi_\theta) := (B(\phi_\theta), A(\phi_\theta)B(\phi_\theta), \dots, A(\phi_\theta)^{n_m-1}B(\phi_\theta)) \quad , \quad O^0(\phi_\theta) := \begin{pmatrix} C(\phi_\theta) \\ C(\phi_\theta)A(\phi_\theta) \\ \vdots \\ C(\phi_\theta)A(\phi_\theta)^{n_m-1} \end{pmatrix}$$

are respectively the controllability ( $C^0$ ) and observability ( $O^0$ ) matrices.

**ASSUMPTION 2 (MINIMALITY)** *The state-space system (10)-(11) is minimal, i.e. such that the rank condition in eq. (12) is valid.*

In general, under Assumptions 1–2, the state-space system (10)-(11) admits a stationary VARMA-type representation for  $y_t$ , see e.g. Hannan and Deistler (1988). The minimality condition mimics the left-coprime condition typically imposed on (or assume in) VARMA processes (see e.g. Lütkepohl, 2005, p. 452). For cases where the  $D(\phi_\theta)$  matrix in eq.s (10)-(11) is square ( $n_y = n_\omega$ ), it is possible to derive conditions that ensure the fundamentalness of the ABCD form (Fernández-Villaverde *et al.* 2007), i.e. a representation where the state  $Z_{m,t}$  is fully recoverable from  $y_1, \dots, y_t$ , and conditions for the existence of a finite-order VAR representation for  $y_t$  (Ravenna, 2007; Franchi and Vidotto, 2013; Franchi and Paruolo, 2015). We do not impose any of these conditions, in the sense that the bootstrap approach we present below, other than not requiring the existence of finite-order VAR representation for  $y_t$ , can also be applied to specifications and adaptations of system (1)-(2) that potentially accommodate foresight phenomena about technology and government spending, anticipated shocks and non-fundamental representations along the lines of e.g. Leeper *et al.* (2013).

Our last assumption regards the (local) identifiability of the structural parameters  $\theta$  from system (10)-(11). Following Komunjer and Ng (2011),  $\theta$  is identified if and only if

$$\text{rank}(\mathcal{M}(\phi_{\theta_0})) = \dim(\theta_0) + n_m^2 + n_\omega^2 \quad (13)$$

where

$$\mathcal{M}(\theta) := \begin{pmatrix} \frac{\partial \text{vec} A(\phi_\theta)}{\partial \theta'} & A(\phi_\theta)' \otimes I_{n_m} - I_{n_m} \otimes A(\phi_\theta) & 0_{n_m^2 \times n_\omega^2} \\ \frac{\partial \text{vec} B(\phi_\theta)}{\partial \theta'} & B(\phi_\theta)' \otimes I_{n_m} & I_{n_\omega} \otimes B(\phi_\theta) \\ \frac{\partial \text{vec} C(\phi_\theta)}{\partial \theta'} & -I_{n_m} \otimes C(\phi_\theta) & 0_{n_y n_m \times n_\omega^2} \\ \frac{\partial \text{vec} D(\phi_\theta)}{\partial \theta'} & 0_{n_y n_m \times n_m^2} & I_{n_\omega} \otimes D(\phi_\theta) \\ \frac{\partial \text{vech} \Sigma_\omega(\phi_\theta)}{\partial \theta'} & 0_{\frac{1}{2}n_y(n_y+1) \times n_m^2} & -2d_{n_\omega}^+(\Sigma_\omega(\phi_\theta) \otimes I_{n_\omega}) \end{pmatrix} .$$

Here  $d_{n_\omega}^+ := (d'_{n_\omega} d_{n_\omega})^{-1} d'_{n_\omega}$  is the Moore-Penrose inverse of the duplication matrix  $d_{n_\omega}$ , i.e. the  $n_\omega^2 \times \frac{1}{2}n_\omega(n_\omega + 1)$  matrix such that  $d_{n_\omega} \text{vech}(\Sigma_\omega) = \text{vec}(\Sigma_\omega)$ , and it is maintained that the matrices  $A(\phi_\theta)$ ,  $B(\phi_\theta)$ ,  $C(\phi_\theta)$  and  $D(\phi_\theta)$  have continuous first and second derivatives with respect to  $\theta$ .

**ASSUMPTION 3 (IDENTIFICATION)** *The state-space system (10)-(11) is ‘locally identified’, i.e. the functions  $\text{vec}(A(g(\theta)))$ ,  $\text{vec}(B(g(\theta)))$ ,  $\text{vec}(C(g(\theta)))$  and  $\text{vec}(D(g(\theta)))$  have continuous first and second derivatives with respect to  $\theta$  and the rank condition in eq. (13) is valid in a neighborhood  $\mathcal{N}_{\theta_0}$  of  $\theta_0$ .*



Komunjer and Ng (2011) show that for the case  $n_y \leq n_\omega$ , the information matrix of the system is nonsingular if and only if the condition in eq. (13) is satisfied. Using the terminology in Canova and Sala (2009), Assumption 3 is a statement about the ‘population’ identification of the DSGE model. However, identification problems in a system of variables featuring highly nonlinear restrictions might also involve the relationship between the structural parameters and the sample objective function, which might display ‘small’ curvature in certain regions of the parameter space. We clarify these issues in the next sections.

Finally, we consider the counterpart of the state space representation in eq.s (10)-(11) which does not incorporate the CER. Indeed, in the next sections a state space representation of the DSGE model which does not incorporate the CER is needed in order to compute QLR tests for the CER.<sup>5</sup> Guerron-Quintana *et al.* (2013) have shown that it is in principle possible to couple the ABCD form with a counterpart where  $\phi_\theta$  is replaced with a vector of parameters, say  $\phi$ , which collapses to  $\phi_\theta$  when the CER are imposed, see also Angelini and Fanelli (2015). We therefore posit that, associated with the ABCD form in eq.s (10)-(11), it exists a state space representation for  $Z_{m,t}$  and  $y_t$ , given by

$$Z_{m,t} = \begin{matrix} A(\phi) & B(\phi) \\ n_m \times n_m & n_m \times n_\omega \end{matrix} \begin{matrix} Z_{m,t-1} \\ \omega_t^u \\ n_m \times 1 & n_\omega \times 1 \end{matrix} + \begin{matrix} \omega_t^u \\ n_\omega \times 1 \end{matrix} \quad (14)$$

$$y_t = \begin{matrix} C(\phi) & D(\phi) \\ n_y \times n_m & n_y \times n_\omega \end{matrix} \begin{matrix} Z_{m,t-1} \\ \omega_t^u \\ n_m \times 1 & n_\omega \times 1 \end{matrix}, \quad t = 1, \dots, T, \quad (15)$$

where  $\omega_t^u$  is an  $n_\omega \times 1$  vector of white noise structural shocks with diagonal covariance matrix  $\Sigma_{\omega^u}$ , and  $\phi$  is a vector of coefficients such that  $\dim(\phi) = \dim(\phi_\theta)$ ,  $\dim(\phi) > \dim(\theta)$ .

ASSUMPTION 4 (STATE SPACE REPRESENTATION WITHOUT THE CER) *System (14)-(15) is such that:*

(4.i) *the matrix  $A(\phi)$  is stable and invertible, and the functions  $\text{vec}(A(\phi))$ ,  $\text{vec}(B(\phi))$ ,  $\text{vec}(C(\phi))$  and  $\text{vec}(D(\phi))$  have continuous first and second derivatives with respect to  $\phi$ ;*

(4.ii) *is in minimal form, in the sense that the condition in eq. (12) is valid once  $\phi_{\theta_0}$  is replaced with  $\phi_0$ , where  $\phi_0$  is the true value of  $\phi$  and is an interior point of the compact parameter space  $\mathcal{P}_\phi$ ;*

(4.iii) *is identified in the sense that the condition in eq. (13) is valid once  $\phi_{\theta_0}$  is replaced with  $\phi_0$ .*

While system (10)-(11) characterizes the first- and second-order moments of the data under the CER, together with the (assumed fixed) initial conditions for the state, system (14)-(15) under Assumption 4.i-4.iii provides a complete characterization of first- and second-order moment properties of the data net of the parametric restrictions stemming

---

<sup>5</sup>Finding an ‘unrestricted’ state space representation of the DSGE model is not a trivial task because of the difficulties associated with the specification of an identified minimal state space form that plays the same role reduced form models have in the context of simultaneous systems of equations, see e.g. Schorfheide (2010), Komunjer and Ng. (2011) and Guerron-Quintana *et al.* (2013).

from the theory. Also in this case, the stochastic process that generates the observables,  $\{y_t\}$ , is covariance stationary.

### 3 TEST OF THE CROSS-EQUATION RESTRICTIONS

The ABCD form in eq.s (10)-(11) and the system (14)-(15) can be regarded as the state space interfaces of the DSGE model. The ABCD form in eq.s (10)-(11) is nested within system (14)-(15), in the sense that system (14)-(15) collapses to the ABCD form in eq.s (10)-(11) once the elements of the vector  $\phi$  are replaced with the elements of the vector  $\phi_\theta = g(\theta)$ , where  $g(\cdot)$  is such that the Jacobian  $(\nabla_\theta \phi_\theta)' \equiv \frac{\partial g(\theta)}{\partial \theta'}$  has full-column rank,  $\dim(\theta)$ , in a neighborhood of  $\theta_0$ . Our Technical Supplement discusses in detail the relationship between  $\phi$ ,  $\phi_\theta$  and  $\theta$ .

We have all the ingredients to compute a QLR test of the CER implied by the DSGE model. In Section 3.1 we briefly address the testing problem and in Section 3.2 we characterize the concepts of strong and weak identification and their consequences on inference.

#### 3.1 TESTING PROBLEM

We consider the testing problem

$$H_0 : \phi = \phi_\theta = g(\theta) \quad \text{against} \quad H_1 : \phi \neq \phi_\theta \quad (16)$$

by a QLR test. The null  $H_0$  incorporates the CER. The state space representation of the DSGE model under  $H_0$  is given by system (10)-(11). Instead, the state space representation of the DSGE model under  $H_1$  corresponds to system (14)-(15). To compute a QLR test of the CER it is necessary to maximize the likelihood associated with both systems. Let  $\ell_T(\phi)$  and  $\ell_T(\phi_\theta) = \ell_T(g(\theta))$  be the log-likelihoods of the DSGE model under  $H_1$  and  $H_0$ , respectively, and let  $\hat{\phi}_T := \arg \max_{\phi \in \mathcal{P}_\phi} \ell_T(\phi)$  and  $\hat{\theta}_T := \arg \max_{\theta \in \mathcal{P}_\theta} \ell_T(\phi_\theta)$  ( $\hat{\phi}_{\theta,T} = g(\hat{\theta}_T)$ ) the QML estimators of  $\phi$  and  $\theta$  ( $\phi_\theta$ ). These estimation problems are based on Gaussian innovation errors and the Kalman-filter and are considered in the Technical Supplement. For the purposes of the present analysis, it is important to recall that under  $H_0$ , the innovation form representation (Anderson and Moore, 1979) in (steady-state form) associated with the ABCD form in eq.s (10)-(11) is summarized by the expressions

$$\hat{Z}_{m,t+1|t} = A(\phi_\theta)\hat{Z}_{m,t|t-1} + K_t(\phi_\theta) \epsilon_t^0 \quad (17)$$

$$y_t = H^+(\phi_\theta)\hat{Z}_{m,t|t-1} + \epsilon_t^0 \quad (18)$$

where  $K_t = K_t(\phi_\theta)$  is the Kalman gain,  $H^+(\phi_\theta) := C(\phi_\theta)A(\phi_\theta)^{-1}$  and  $\hat{Z}_{m,t|t-1} := E(Z_{m,t} | \mathcal{F}_{t-1}^y)$ , where  $\mathcal{F}_t^y := \sigma(y_t, \dots, y_1) \subset \mathcal{F}_t$  is the information set based on the observable variables up to time  $t$ .  $\epsilon_t^0$  are the innovation errors with covariance matrix  $\Sigma_{\epsilon^0,t}$ ,  $t = 1, \dots, T$ . We use the superscript '0' for  $\epsilon_t^0$  and  $\Sigma_{\epsilon^0,t}$  to remark that the representation in eq.s (17)-(18) is obtained under the null  $H_0$  which imposes the CER. Under the alternative  $H_1$ , the innovation form representation is similar to system (17)-(18) but  $\phi_\theta$  needs to be replaced with  $\phi$ , hence  $\epsilon_t^u := y_t - H^+(\phi)\hat{Z}_{m,t|t-1}$ , and the implied covariance matrix is  $\Sigma_{\epsilon^u,t}$ ,  $t = 1, \dots, T$ .

The QLR test for the CER is given by

$$QLR_T(\hat{\theta}_T) := -2[\ell_T(\hat{\phi}_{\theta,T}) - \ell_T(\hat{\phi}_T)]. \quad (19)$$

The asymptotic properties of the tests statistics  $QLR_T(\hat{\theta}_T)$  are intimately related to the asymptotic properties of  $\hat{\theta}_T$  ( $\hat{\phi}_{\theta,T}$ ) and  $\hat{\phi}_T$  and these crucially depend on whether the regularity conditions for inference are valid in the estimated DSGE model.

### 3.2 STRONG AND WEAK IDENTIFICATION: CHARACTERIZATION

While the DSGE model is unidentified if the rank condition in Assumption 3 fails, the validity of Assumption 3 does not necessarily rule out cases in which the log-likelihood of the system  $\ell_T(g(\theta))$  does not satisfy the regularity conditions permitting ‘standard inference’, where by this term we denote the situation where  $\hat{\theta}_T$  is consistent and asymptotically Gaussian, with quadratic forms derived from  $\hat{\theta}_T$  being asymptotically  $\chi^2$ -distributed. More precisely, the DSGE model can satisfy the local identification condition in Assumption 3 at  $\theta = \theta_0$ , but  $\ell_T(g(\theta))$  might exhibit local maxima (and minima) and nearly flat surfaces in some directions of the parameter space. This concept is often termed, borrowing terminology from the literature on instrumental variables and generalized methods of moments, ‘weak identification’ (Staiger and Stock, 1997; and Stock and Wright, 2000). The same might occur to  $\ell_{o,T}(\phi)$ .

When the standard regularity conditions for inference hold (see the Technical Supplement, Section TS.4.2), we can think about the scenario in which Assumptions 1-3 are also valid for a  $\theta^\dagger$  in  $\mathcal{P}^D$  that replaces  $\theta_0$  and such that

$$\hat{\theta}_T \xrightarrow{p} \theta^\dagger \quad (20)$$

where  $\theta^\dagger = \theta_0$  under  $H_0$ ; furthermore, and

$$T^{1/2} Z_T \xrightarrow{d} N(0_{\dim(\theta) \times 1}, I_{\dim(\theta)}), \quad (21)$$

where  $T^{1/2} Z_T := T^{1/2} \check{\mathcal{I}}_{\hat{\theta}_T, T}^{1/2}(\hat{\theta}_T - \theta^\dagger)$ . The quantity  $\check{\mathcal{I}}_{\hat{\theta}_T, T}$  in  $Z_T$  is a consistent, (a.s.) positive definite estimator of the inverse of the covariance matrix

$$V_{\theta^\dagger} := \left( \mathcal{I}_{\theta^\dagger, \infty}^{2D} \left( \mathcal{I}_{\theta^\dagger, \infty}^{OP} \right)^{-1} \mathcal{I}_{\theta^\dagger, \infty}^{2D} \right)^{-1} \quad (22)$$

where, with  $\nabla_{\theta} \ell_{o,T}(\theta^\dagger) := \frac{\partial \ell_{o,T}(\theta)}{\partial \theta} \Big|_{\theta=\theta^\dagger}$  and  $\nabla_{\theta\theta}^2 \ell_{o,T}(\theta^\dagger) := \frac{\partial^2 \ell_{o,T}(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\theta^\dagger}$ ,

$$\mathcal{I}_{\theta^\dagger, \infty}^{OP} := \lim_{T \rightarrow \infty} \frac{1}{T} \mathcal{I}_{\theta^\dagger, T}^{OP}, \quad \mathcal{I}_{\theta^\dagger, T}^{OP} := E \left( \nabla_{\theta} \ell_{o,T}(\theta^\dagger) \times \nabla_{\theta} \ell_{o,T}(\theta^\dagger)' \right) \quad (23)$$

$$\mathcal{I}_{\theta^\dagger, \infty}^{2D} := \lim_{T \rightarrow \infty} \frac{1}{T} \mathcal{I}_{\theta^\dagger, T}^{2D}, \quad \mathcal{I}_{\theta^\dagger, T}^{2D} := -E \left( \nabla_{\theta\theta}^2 \ell_{o,T}(\theta^\dagger) \right). \quad (24)$$

The (locally) identified parameter vector  $\theta^\dagger$  plays in eq.s (20)-(22) the role of ‘pseudo-true’ value of the structural parameters which minimizes the Kullback-Leibler information criterion (White, 1982). Under standard regularity conditions, the QML estimator

$\hat{\theta}_T$  converges to  $\theta^\dagger$  when the CER are imposed in estimation. When  $H_0$  in eq. (16) is valid,  $\theta^\dagger = \theta_0$  in eq.s (20)-(22) and  $V_{\theta^\dagger} = V_{\theta_0}$  reads as the asymptotic covariance matrix of  $T^{1/2}(\hat{\theta}_T - \theta_0)$ , i.e.  $V_{\theta_0} := \lim_{T \rightarrow \infty} \text{Var}(T^{1/2}\hat{\theta}_T)$ . Furthermore, if the innovation errors that characterize the Kalman filter algorithm are Gaussian and  $H_0$  holds, both matrices  $\tilde{\mathcal{I}}_{\theta_0, T}^{OP}$  and  $\tilde{\mathcal{I}}_{\theta_0, T}^{2D}$  derived from eq.s (23)-(24) for  $\theta^\dagger = \theta_0$  characterize the system's information matrix and  $V_{\theta_0}$  collapses to the expression  $V_{\theta_0} := (\tilde{\mathcal{I}}_{\theta_0, \infty}^{2D})^{-1}$ . The quantity:

$$\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{OP} := \frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \ell_t(\hat{\theta}_T) \times \nabla_{\theta} \ell_t(\hat{\theta}_T)', \quad (25)$$

known as ‘incremental observed information’ and the quantity:

$$\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{2D} := -\frac{1}{T} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t(\hat{\theta}_T), \quad (26)$$

known as ‘observed information’ (evaluated at  $\hat{\theta}_T$ ), are both consistent (interchangeable) estimates of the system's information matrix. It turns out that one can use

$$\tilde{\mathcal{I}}_{\hat{\theta}_T, T} = \tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{2D} \left( \tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{OP} \right)^{-1} \tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{2D} \quad (27)$$

for consistent estimation of the information matrix under  $H_0$ .

Similar considerations apply to the estimation of the parameters in  $\phi$  under Assumption 4 and  $H_1$ , see Watson (1989). It follows that under Assumptions 1-4 (and the regularity conditions (TS-A.i)-(TS-A.viii) reviewed in the Technical Supplement), standard arguments imply that the test statistic  $QLR_T(\hat{\theta}_T)$  in eq. (19) is asymptotically  $\chi^2$ -distributed with  $k = \dim(\phi) - \dim(\theta)$  degree of freedom (henceforth  $\chi_k^2$ ) when  $H_0$  is true and diverges under  $H_1$ . Thus, the null  $H_0$  is rejected when  $QLR_T(\hat{\theta}_T) \geq c_{\chi_k^2}^\eta$ , where  $c_{\chi_k^2}^\eta$  is the  $\eta$ -level cut-off point associated with the  $\chi_k^2$ -distribution and  $0 < \eta < 1$  is the pre-fixed nominal significance level (or type-I error probability) of the test. Notice, however, that also in strongly identified DSGE models, the  $\chi_k^2$  distribution might provide a poor approximation in finite samples.

Conversely, when standard regularity conditions do not hold as e.g. in weakly identified DSGE models, the consistency result in eq. (20) and the asymptotic normality result in eq. (21) are no longer guaranteed, see, among others, Andrews and Cheng (2012) and Andrews and Mikusheva (2015). Andrews and Mikusheva (2015) show through examples that under weak identification, the (appropriately normalized) quadratic variation of the score converges to a fixed positive definite matrix, while the Hessian converges in distribution to a random matrix. Thus, under weak identification, one can expect large disparities between different estimators of information also if the model respects White's (1982) information matrix equality. Also in the case where the innovations errors are Gaussian, it is not necessarily true that the matrices in eq.s (25) and (26) evaluated at  $\theta = \theta_0$  are consistent (and interchangeable) estimates of the information matrix. Andrews and Mikusheva (2015) observe that in these cases ‘it is unwise to estimate the information matrix using an estimator of  $\theta$ ’ such as e.g. the quantities  $\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{2D}$  and  $\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{OP}$  in eq.s (25)-(26).

Summing up, the class of weakly identified DSGE models we have in mind in this paper, must not be considered misspecified in the ‘conventional’ sense discussed in e.g. Inoue *et al.* (2015), An and Shorfheide (2007, Section 3.1) and Angelini and Fanelli (2015). Rather, weak identification can be associated with situations in which the convergences in eq.s (20)-(21) are no longer valid and suitably normalized versions of the matrix  $\tilde{Z}_{\theta_0, T}^{2D}$  converge weakly to a random matrix. Conventionally, throughout the paper we use the terms ‘strong identification’ and ‘strongly identified DSGE model’ to denote the situations in which the regularity conditions that permit standard inference on  $\theta$  are valid. Conversely, failure of these conditions, as it happens e.g. in ‘weakly identified’ DSGE models, does not permit the ‘conventional’ asymptotic expansions of the likelihood, with the consequence that irrespective of whether the CER are true (i.e.  $H_0$  is valid) or not (i.e.  $H_1$  is valid), the test statistics  $T^{1/2}Z_T$  deviates asymptotically from the Gaussian distribution.

The different asymptotic behaviour of the statistic  $T^{1/2}Z_T$  under strong identification and when strong identification fails is at the basis of our bootstrap approach to the evaluation of DSGE models.

## 4 BOOTSTRAP APPROACH

We take the point of view of a practitioner who is concerned with (i) the estimation of the structural parameters of the DSGE model in eq.s (1)-(2) and (ii) the empirical evaluation of the estimated model by the testing the hypothesis  $H_0$  against  $H_1$  in eq.(16). We assume that, given the estimation sample  $y_1, \dots, y_T$ , estimation of  $\theta$  ( $\phi_\theta$ ) (under  $H_0$ ) and  $\phi$  (under  $H_1$ ) has been completed and the QML estimates  $\hat{\theta}_T$  ( $\hat{\phi}_{\theta, T} = g(\hat{\theta}_T)$ ) and  $\hat{\phi}_T$  are available. The practitioner does not know *a priori* whether the DSGE model meets all regularity conditions that permit standard inference.

Our bootstrap approach is based on the following idea. After the QLR test in eq. (19) is computed on the sample  $y_1, \dots, y_T$ , we construct its bootstrap analog, denoted by  $QLR_T^*(\hat{\theta}_T^*)$ . This can be done by a standard bootstrap algorithm as the one described in Section 4.1 below. Henceforth, with ‘\*’ we denote bootstrap analogs of estimators and statistics; for instance,  $\hat{\theta}_T^*$  denotes the estimator of  $\theta$  obtained on the bootstrap sample  $y_1^*, \dots, y_T^*$  (the construction of the bootstrap sample is detailed below). In order to compute the bootstrap p-value associated to the QLR test, it is required to generate an arbitrary number of bootstrap samples,  $B$  say, under the null of the DSGE model, i.e. with the CER in  $H_0$  imposed (that is, with  $\theta$  fixed at the QML estimate  $\hat{\theta}_T$ ). Then, on each generated bootstrap sample, the state space representation of the model is estimated both under the CER ( $H_0$ ) and unrestrictedly ( $H_1$ ). In Section 4.2 we prove that in strongly identified DSGE models the bootstrap is consistent. Thus, conditional on the original data,  $\hat{\theta}_T^*$  should be normally distributed for large values of  $T$  and  $B$ . Accordingly, each element of  $\hat{\theta}_T^*$  is expected (conditional on the original data) to conform to the Gaussian distribution in large samples. In DSGE models where the regularity condition for standard inference fail, instead,  $\hat{\theta}_T^*$  may not conform to the Gaussian distribution. In Section 4.3 we discuss how these considerations can be used to design a bootstrap-based misspecification test for the estimated DSGE model. In Section 4.4

we compare our approach with the existing literature.

#### 4.1 BOOTSTRAP ALGORITHM

We employ a nonparametric ‘restricted bootstrap’ algorithm (see e.g. Davidson and MacKinnon, 1999), where the bootstrap samples are generated using the parameter estimates  $\hat{\theta}_T$  obtained under  $H_0$ . The QLR test statistic,  $QLR_T(\hat{\theta}_T)$ , computed as in eq. (19) is stored, along with  $\hat{\theta}_T$ . Our procedure is adapted from Stoffer and Wall (1991), see also Cavanaugh and Shumway (1997), and is described through the following algorithm. Here, steps 1–4 define the bootstrap sample, the bootstrap parameter estimators and related bootstrap QLR statistic; steps 5–7 describe the numerical computation of the bootstrap p-value associated to the bootstrap QLR test.<sup>6</sup>

##### ALGORITHM (RESTRICTED BOOTSTRAP)

1. Given the innovation residuals  $\hat{\epsilon}_t^0 := y_t - H^+(\hat{\phi}_{\theta,T})\hat{Z}_{m,t|t-1}$  and the estimated covariance matrices  $\hat{\Sigma}_{\epsilon^0,t} = \Sigma_{\epsilon^0,t}(\hat{\phi}_{\theta,T})$  produced by the estimation of the DSGE model under the CER ( $H_0$ ), construct the standardized innovations as

$$\hat{e}_t^0 := \hat{\Sigma}_{\epsilon^0,t}^{-1/2} \hat{\epsilon}_t^{0,c}, \quad t = 1, \dots, T, \quad (28)$$

where  $\hat{\Sigma}_{\epsilon^0,t}^{-1/2}$  is the inverse of the square-root matrix of  $\hat{\Sigma}_{\epsilon^0,t}$  and  $\hat{\epsilon}_t^{0,c}$ ,  $t = 1, \dots, T$ , are the centered residuals  $\hat{\epsilon}_t^{0,c} := \hat{\epsilon}_t^0 - T^{-1} \sum_{t=1}^T \hat{\epsilon}_t^0$ ;

2. Sample, with replacement,  $T$  times from  $\hat{e}_1^0, \hat{e}_2^0, \dots, \hat{e}_T^0$  to obtain the bootstrap sample of standardized innovations  $e_1^*, e_2^*, \dots, e_T^*$ ;
3. Mimicking the innovation form representation of the DSGE model in eq.s (17)-(18), the bootstrap sample  $y_1^*, y_2^*, \dots, y_T^*$  is generated recursively by solving, for  $t = 1, \dots, T$ , the system

$$\begin{pmatrix} \hat{Z}_{m,t+1|t}^* \\ y_t^* \end{pmatrix} = \begin{pmatrix} A(\hat{\phi}_{\theta,T}) & 0_{n_m \times n_y} \\ H^+(\hat{\phi}_{\theta,T}) & 0_{n_y \times n_y} \end{pmatrix} \begin{pmatrix} \hat{Z}_{m,t|t-1}^* \\ y_{t-1}^* \end{pmatrix} + \begin{pmatrix} K_t(\hat{\phi}_{\theta,T})\hat{\Sigma}_{\epsilon^0,t}^{1/2} \\ \hat{\Sigma}_{\epsilon^0,t}^{1/2} \end{pmatrix} e_t^* \quad (29)$$

with initial condition  $\hat{Z}_{m,1|0}^* = \hat{Z}_{m,1|0}$ ;

- 4 From the generated pseudo-sample  $y_1^*, y_2^*, \dots, y_T^*$ , estimate the DSGE model under  $H_0$  obtaining the bootstrap estimator  $\hat{\theta}_T^*$  ( $\hat{\phi}_{\theta,T}^* = g(\hat{\theta}_T^*)$ ) and the associated log-likelihood  $\ell_T^*(\hat{\phi}_{\theta,T}^*)$ , and estimate the DSGE model under  $H_1$  obtaining the bootstrap estimator  $\hat{\phi}_T^*$  and the associated log-likelihood  $\ell_T^*(\hat{\phi}_T^*)$ ; the bootstrap QLR test for the CER is defined as:

$$QLR_T^*(\hat{\theta}_T^*) := -2[\ell_T^*(\hat{\phi}_{\theta,T}^*) - \ell_T^*(\hat{\phi}_T^*)]; \quad (30)$$

<sup>6</sup>Matlab codes used in the simulation study (Section 5), in the empirical illustration (Section 6) and in the Technical Supplement are available upon request.

5 Steps 2-4 are repeated  $B$  times in order to obtain  $B$  bootstrap realizations of  $\hat{\theta}_T$  and  $\hat{\phi}_T$ , say  $\{\hat{\theta}_{T:1}^*, \hat{\theta}_{T:2}^*, \dots, \hat{\theta}_{T:B}^*\}$  and  $\{\hat{\phi}_{T:1}^*, \hat{\phi}_{T:2}^*, \dots, \hat{\phi}_{T:B}^*\}$ , and the  $B$  bootstrap realizations of the associated bootstrap QLR test,  $\{QLR_{T:1}^*, QLR_{T:2}^*, \dots, QLR_{T:B}^*\}$ , where  $QLR_{T:b}^* := QLR_T^*(\hat{\theta}_{T:b}^*)$ ,  $b = 1, \dots, B$ ;

6 The bootstrap  $p$ -value of the test of the CER is computed as

$$\hat{p}_{T,B}^* := \hat{G}_{T,B}^*(QLR_T(\hat{\theta}_T)) \quad , \quad \hat{G}_{T,B}^*(x) := B^{-1} \sum_{b=1}^B \mathbb{I}\{QLR_{T:b}^* > x\}, \quad (31)$$

$\mathbb{I}\{\cdot\}$  being the indicator function;

7 The bootstrap QLR test for the CER at the  $100\eta\%$  (nominal) significance level rejects  $H_0$  if  $\hat{p}_{T,B}^* \leq \eta$ .

Some remarks are in order.

REMARK 4.1 Although the bootstrap algorithm features a ‘restricted’ resampling scheme, the procedure can be easily adapted to the case of an ‘unrestricted’ resampling scheme, where the bootstrap samples are generated using unrestricted parameter estimates, i.e. without imposing the CER. In this case, in step 1 one may consider and resample the residuals and associated covariance matrices obtained from the estimation of the state space model under  $H_1$ , i.e.  $\hat{\epsilon}_t^u := y_t - H^+(\hat{\phi}_T)\hat{Z}_{m,t|t-1}$  and  $\hat{\Sigma}^{\epsilon^u,t} = \Sigma^{\epsilon^u,t}(\hat{\phi}_T)$ , and then replace  $\hat{\phi}_{\theta,T}$  with  $\hat{\phi}_T$  in step 3.

REMARK 4.2 The algorithm is ‘nonparametric’ in the sense that in step 1 the bootstrap innovations are obtained as random draws from the standardized residuals  $\hat{\epsilon}_t^0$ ,  $t = 1, \dots, T$ ; that is, the so-called nonparametric ‘i.i.d. bootstrap’ is used to generate the innovations. However, if the normality hypothesis holds true, one may alternatively employ a parametric version of the bootstrap algorithm, which simply requires ignoring steps 1 and 2 and starting from the step 3, with the  $e_t^*$  now taken as independent random draws from the  $N(0_{n_y \times 1}, I_{n_y})$  distribution.

REMARK 4.3 An alternative approach to our i.i.d. bootstrap is the so-called ‘wild bootstrap’ (Wu, 1986; Liu, 1988; Mammen, 1992), which allows to mimick possible (conditional and unconditional) heteroskedasticity patterns present in the the original data; see e.g. Goncalves and Kilian (2004) for the case of stationary autoregressions and Boswijk et al. (2016) for hypothesis testing in multivariate models. In our framework, the wild bootstrap shocks would be generated in Step 3 as

$$e_{w,t}^* := \hat{\epsilon}_t^0 w_t^*, \quad t = 1, \dots, T$$

where  $w_t^*$  is an i.i.d. scalar sequence with 0 mean, unit variance, and finite fourth order moments. Using the wild bootstrap, the standardization of the residuals in (28) (Step 1) is no longer necessary (since, conditionally on the original data,  $E^*(e_{w,t}^* e_{w,t}^{*'}) = \hat{\epsilon}_t^0 \hat{\epsilon}_t^{0'}$ ), and consequently the recursion in (29) can be replaced by the simpler recursion

$$\begin{pmatrix} \hat{Z}_{m,t+1|t}^* \\ y_t^* \end{pmatrix} = \begin{pmatrix} A(\hat{\phi}_{\theta,T}) & 0_{n_m \times n_y} \\ H^+(\hat{\phi}_{\theta,T}) & 0_{n_y \times n_y} \end{pmatrix} \begin{pmatrix} \hat{Z}_{m,t|t-1}^* \\ y_{t-1}^* \end{pmatrix} + \begin{pmatrix} K_t(\hat{\phi}_{\theta,T}) \\ I_{n_y} \end{pmatrix} e_{w,t}^*,$$

for  $t = 1, \dots, T$ , again initialized at  $\hat{Z}_{m,1|0}^* = \hat{Z}_{m,1|0}$ .

REMARK 4.4 In the restrictive case where  $D(\phi_\theta)$  in eq.s (10)-(11) is a square ( $n_y = n_\omega$ ) and invertible matrix, a finite-order VAR representation for  $y_t$  exists (see Section 2). In this case, the bootstrap algorithm can be adapted such that the VAR representation of the DSGE model is used to estimate the parameters and generate the residuals used in the resampling scheme, see e.g. Bårdsen and Fanelli (2005) and Cho and Moreno (2006).

## 4.2 ASYMPTOTIC VALIDITY

Assume that  $\mathbf{H}_0$  in eq.(16) holds and  $\theta^\dagger = \theta_0$ . Let  $T^{1/2}Z_T^* := T^{1/2}\check{\mathcal{I}}_{\hat{\theta}_T, T}^{1/2}(\hat{\theta}_T^* - \hat{\theta}_T)$  be the bootstrap analog of  $T^{1/2}Z_T := T^{1/2}\check{\mathcal{I}}_{\hat{\theta}_T, T}^{1/2}(\hat{\theta}_T - \theta_0)$ , where  $\check{\mathcal{I}}_{\hat{\theta}_T, T}$  is a consistent estimator of the inverse of the asymptotic covariance matrix  $V_{\theta_0}$ , see eq. (27). The QML estimator  $\hat{\theta}_T$ , used to generate the bootstrap sample, plays in  $T^{1/2}Z_T^*$  the role of a ‘pseudo-true parameter’. Random draws from  $\hat{\theta}_T^*$  can be obtained through the algorithm presented in previous section, which generates  $B$  bootstrap realizations  $\hat{\theta}_{T:1}^*, \dots, \hat{\theta}_{T:B}^*$ . With respect to estimation under  $\mathbf{H}_1$ , we have  $T^{1/2}U_T := T^{1/2}\check{\mathcal{I}}_{\hat{\phi}_T, T}^{1/2}(\hat{\phi}_T - \phi_0)$  and its bootstrap analog  $T^{1/2}U_T^* := T^{1/2}\check{\mathcal{I}}_{\hat{\phi}_T, T}^{1/2}(\hat{\phi}_T^* - \hat{\phi}_T)$ .

Using the regularity conditions by Ljung and Caines (1979), Stoffer and Wall (1991) prove that their bootstrap algorithm ensures that (a)  $T^{1/2}U_T^*$  has the same asymptotic distribution of  $T^{1/2}U_T$ , i.e. multivariate Gaussian and (b) the bootstrap standard errors are consistent. Stoffer and Wall (1991) do not consider any constrained estimation problem (hence they only focus on the case  $\mathbf{H}_1$ ). In the framework of the DSGE models considered here,  $\phi = \phi_\theta = g(\theta)$  under  $\mathbf{H}_0$ , and Stoffer and Wall’s (1991) result (a) can be restated in the following form

$$\hat{\theta}_T^* - \hat{\theta}_T \xrightarrow{P^*} 0_{\dim(\theta) \times 1} \quad (32)$$

$$T^{1/2}Z_T^* \xrightarrow{d^*} N(0_{\dim(\theta) \times 1}, I_{\dim(\theta)}) \quad (33)$$

where ‘ $\xrightarrow{P^*}$ ’ denotes convergence ‘in  $P^*$ , in probability’, while ‘ $\xrightarrow{d^*}$ ’ denotes converge in ‘conditional distribution, in probability’ (see Appendix A). Stoffer and Wall’s (1991) result (b) can be stated by:

$$\text{Var}^*(T^{1/2}\hat{\theta}_T^*) - \text{Var}(T^{1/2}\hat{\theta}_T) \xrightarrow{P^*} 0_{\dim(\theta) \times \dim(\theta)} \quad (34)$$

where  $\text{Var}^*(\cdot)$  denotes variance conditional on the original sample (see Appendix A.1). Note that an arbitrarily accurate estimate of  $\text{Var}^*(\hat{\theta}_T^*)$  can be obtained, for sufficiently large  $B$ , from the bootstrap realizations  $\hat{\theta}_{T:1}^*, \hat{\theta}_{T:2}^*, \dots, \hat{\theta}_{T:B}^*$  as

$$\widehat{\text{Var}}^*(\hat{\theta}_T^*) := \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_{T:b}^* - \bar{\theta}_T^*)(\hat{\theta}_{T:b}^* - \bar{\theta}_T^*)' \quad , \quad \bar{\theta}_T^* := \frac{1}{B} \sum_{b=1}^B \hat{\theta}_{T:b}^*$$



The square root of the elements on the main diagonal of  $\widehat{\text{Var}}^*(\hat{\theta}_T^*)$  delivers the ‘bootstrap standard errors’ routinely reported by practitioners along with the QML estimates of the structural parameters.

We can now establish that the bootstrap is indeed consistent in the framework of the DSGE models considered in this paper. This result is given in the next proposition, where it is shown that the convergences in eq.s (32)-(34) hold in strongly identified DSGE models and, in addition, that the bootstrap is consistent. The following notation is used in the proposition: for  $x := (x_1, \dots, x_{\dim(\theta)})' \in \mathbb{R}^{\dim(\theta)}$  and  $K_T := T^{1/2}Z_T = (K_{1,T}, \dots, K_{\dim(\theta),T})'$ ,  $G_T(x) := P(K_{1,T} \leq x_1, \dots, K_{\dim(\theta),T} \leq x_{\dim(\theta)})$  denotes the cumulative distribution function (CDF) of  $K_T$ ; moreover, for  $K_T^* := T^{1/2}Z_T^* = (K_{1,T}^*, \dots, K_{\dim(\theta),T}^*)'$ ,  $G_T^*(x) := P^*(K_{1,T}^* \leq x_1, \dots, K_{\dim(\theta),T}^* \leq x_{\dim(\theta)})$  denotes the CDF of  $K_T^*$ , conditional on the original sample.

**PROPOSITION 1 (CONSISTENCY)** *Consider the DSGE model in eq.s (1)-(2) with ABCD representation in eq.s (10)-(11), Assumptions 1-3, the set of regularity conditions (A.i\*)-(A.viii\*) in the Appendix. Assume further that the CER are true, i.e.  $H_0$  is valid and  $\theta^\dagger = \theta_0$ . Then the following results holds as  $T \rightarrow \infty$ :*

- (i)  $\hat{\theta}_T^*$  satisfies the convergences in eq.s (32)-(34);
- (ii)

$$\sup_{x \in \mathbb{R}^{\dim(\theta)}} |G_T^*(x) - G_\infty(x)| \xrightarrow{p} 0 \quad (35)$$

where  $G_\infty(x) = \Phi(x)$  and  $\Phi(x)$  is the CDF of the standardized multivariate Gaussian distribution.

**PROOF:** See Appendix A. ■

Some remarks are in order.

**REMARK 4.5** Proposition 1 generalizes Stoffer and Wall’s (1991) main result on state space models along two dimensions. First, it addresses issues related to a constrained (under  $H_0$ ) estimation problem. Second, other than formalizing the fact that the asymptotic distribution of  $T^{1/2}Z_T^* := T^{1/2}\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{1/2}(\hat{\theta}_T^* - \hat{\theta}_T)$  is multivariate Gaussian as the asymptotic distribution of  $T^{1/2}Z_T := T^{1/2}\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{1/2}(\hat{\theta}_T - \theta_0)$  and that the bootstrap standard errors consistently estimate the analytic standard errors of the QML estimator, it further proves the consistency of the bootstrap. Thus, Proposition 1 ensures that one can build and interpret the bootstrap confidence intervals for the structural parameters and for the impulse responses in the ‘conventional’ way.

**REMARK 4.6** By the same arguments, it is possible to prove that under the analog of the regularity conditions of Proposition 1, also the statistic  $T^{1/2}U_T^* := T^{1/2}\tilde{\mathcal{I}}_{\hat{\phi}_T, T}^{1/2}(\hat{\phi}_T^* - \hat{\phi}_T)$ , which is the bootstrap analog of  $T^{1/2}U_T := T^{1/2}\tilde{\mathcal{I}}_{\hat{\phi}_T, T}^{1/2}(\hat{\phi}_T - \phi_0)$ , is asymptotically multivariate Gaussian in strongly identified DSGE models. Therefore, by using standard arguments it can be shown that in strongly identified DSGE models, the asymptotic distribution of the bootstrap QLR test is  $\chi_k^2$  under the null  $H_0$ .

In the next section we show how the consistency result of Proposition 1 can be exploited to design a novel, computational straightforward, bootstrap-based misspecification test for an estimated DSGE model.

### 4.3 A NOVEL BOOTSTRAP MISSPECIFICATION TEST

Other than producing the bootstrap p-value for the QLR test of the CER (as well as standard errors for the estimated parameters), the bootstrap algorithm presented in section 4.1 can also play a relevant role in assessing how ‘close or distant’ is the estimated DSGE model from the case where all regularity conditions for QML-based inference hold. In order to do so, we propose an econometric framework based on the sampling distribution of the bootstrap statistics.

Specifically, consider a bootstrap statistic of the form  $K_T^* := T^{1/2}Z_T^*$ , temporarily assumed to be scalar, and assume that under the stated regularity conditions  $K_T^*$  converges to a random variable  $Z$ , whose CDF is denoted by  $\Phi_Z(\cdot)$ . That is, the CDF of  $K_T^*$ , say  $G_T^*(\cdot)$ , satisfies  $G_T^*(x) - \Phi_Z(x) \rightarrow_p 0$  uniformly in  $x$  as  $T \rightarrow \infty$ ; see Proposition 1 above. Let  $\tilde{B} \leq B$  denote the number of bootstrap repetitions we use to build our misspecification test, where  $B$  is the total (arbitrarily large) number of bootstrap repetitions exploited in the bootstrap algorithm of Section 4.1. We aim at evaluating the regularity conditions underlying the DSGE model by testing whether the deviation of the distribution of  $K_T^*$  - as estimated over the set of  $\tilde{B}$  bootstrap replications - from its theoretical asymptotic limit  $\Phi_Z$  are large enough so that the claim that  $K_T^*$  is asymptotically distributed as  $Z$  may reasonably be rejected. If  $Z$  is the Gaussian, this can be done, under some appropriate conditions, by applying a simple normality test on the  $\tilde{B}$  bootstrap realizations  $K_{T:1}^*, \dots, K_{T:\tilde{B}}^*$ .

It is important to stress that for  $T$  fixed, the distribution of  $K_T^*$  (conditional on the original data) will in general deviate from the normal, even in cases where asymptotic normality holds. Therefore, our idea is to evaluate whether such deviations are large enough such that the hypothesis of *asymptotic* normality of the bootstrap statistic can safely be rejected. If this is the case, then one could interpret the result as signalling a misspecification of the model. For the rest of the discussion, we assume that  $Z$  is standard normal (and  $\Phi_Z$  is therefore the CDF of a  $N(0, 1)$  random variable), although the reasoning may be extended to any given distribution.<sup>7</sup>

Given  $\tilde{B} \leq B$ , we can estimate  $G_T^*(x)$  (assumed to be strictly positive) by the empirical distribution function of  $K_{1:1}^*, \dots, K_{1:\tilde{B}}^*$  as

$$G_{T,\tilde{B}}^*(x) := \tilde{B}^{-1} \sum_{b=1}^{\tilde{B}} \mathbb{I}\{K_{T:b}^* \leq x\}$$

where  $K_{T:b}^*$ ,  $b = 1, \dots, \tilde{B}$ , are i.i.d. draws from  $G_T^*(x)$ , the conditional distribution of  $K_T^*$ , given the original sample. For any fixed  $x \in \mathbb{R}$ , it holds that, as  $\tilde{B} \rightarrow \infty$ ,  $G_{T,\tilde{B}}^*(x) \rightarrow G_T^*(x)$  (almost surely) and, by the CLT,

$$\tilde{B}^{1/2} V_T(x)^{-1/2} (G_{T,\tilde{B}}^*(x) - G_T^*(x)) \xrightarrow{d} N(0, 1) \quad (36)$$

<sup>7</sup>For instance, one could potentially apply this approach to the (asymptotically  $\chi^2$ ) bootstrap QRL test statistic for the CER.

where  $V_T(x) := G_T^*(x)(1 - G_T^*(x))$  can consistently be estimated (as  $\tilde{B} \rightarrow \infty$ ) by  $V_{T,\tilde{B}}(x) := G_{T,\tilde{B}}^*(x)(1 - G_{T,\tilde{B}}^*(x))$ ; see e.g. Shorack and Wellner (1986).

Consider now the statistic  $SN_{T,\tilde{B}}(x) := \tilde{B}^{1/2}V_{T,\tilde{B}}(x)^{-1/2}(G_{T,\tilde{B}}^*(x) - \Phi_Z(x))$ , i.e. the (normalized) distance between the estimated (over  $\tilde{B}$  repetitions) bootstrap distribution  $G_{T,\tilde{B}}^*(x)$  and its theoretical *asymptotic* distribution arising when standard regularity conditions hold.<sup>8</sup> We can decompose  $SN_{T,\tilde{B}}(x)$  as follows:

$$SN_{T,\tilde{B}}(x) = \tilde{B}^{1/2}V_{T,\tilde{B}}(x)^{-1/2}(G_{T,\tilde{B}}^*(x) - G_T^*(x)) + \tilde{B}^{1/2}V_{T,\tilde{B}}(x)^{-1/2}(G_T^*(x) - \Phi_Z(x)). \quad (37)$$

By the CLT in eq. (36), the first term on the right-hand side of eq. (37) will always converge, as  $\tilde{B} \rightarrow \infty$ , to a  $N(0, 1)$  variate. Regarding the leading factor of the second term,  $G_T^*(x) - \Phi(x)$ , under our regularity conditions it is reasonable to assume that it admits a standard Edgeworth expansion, such that  $G_T^*(x) - \Phi(x) = O_p(T^{1/2})$ . This implies that for any  $x$ , the term  $T^{1/2}V_{T,\tilde{B}}(x)^{-1/2}(G_T^*(x) - \Phi_Z(x))$  in the right-hand side of eq. (37) is of  $O_p(1)$ . Therefore,  $\tilde{B}^{1/2}V_{T,\tilde{B}}(x)^{-1/2}(G_T^*(x) - \Phi(x)) = O_p(\tilde{B}^{1/2}T^{-1/2})$  and converges to zero provided  $\tilde{B} = o(T)$  as both  $B$  and  $T$  diverges. That is,  $SN_{\tilde{B}}(x)$  is asymptotically  $N(0, 1)$  under the condition

$$\tilde{B}^{-1} + T^{-1} + \tilde{B}T^{-1} \rightarrow 0. \quad (38)$$

Conversely, if the consistency result of Proposition 1 does not hold and  $G_T(x)$  does not converge (in probability) to  $\Phi_Z(x)$ , then the second term on the right hand side of eq. (37) does not vanishes asymptotically and  $SN_{\tilde{B}}(x)$  diverges at the rate of  $\tilde{B}$ , leading to a rejection of the hypothesis that  $G_T^*(x)$  converges to  $\Phi_Z(x)$  as  $T \rightarrow \infty$ . This is for instance what we expect to happen in weakly identified DSGE models. Obviously, these arguments can easily be extended to the case where  $K_T^* := T^{1/2}Z_T^*$  is a  $\dim(\theta) \times 1$  vector.

The asymptotic behaviour of  $SN_{\tilde{B}}(x)$  under strong identification as well as under cases where the regularity conditions for standard inference fail, suggests a computational straightforward bootstrap-based misspecification test for DSGE models. The test can be constructed as follows. For a proper choice of  $\tilde{B}$  (which is takes as small relatively to  $T$ , see the condition in (38)) and for any  $i = 1, \dots, \dim(\theta)$ , we can apply standard normality tests to the sequence of bootstrap repetitions  $\hat{\theta}_{i,T:1}^*, \hat{\theta}_{i,T:2}^*, \dots, \hat{\theta}_{i,T:\tilde{B}}^*$  obtained using the bootstrap algorithm of Section 4.1. The same can be done, for any  $j = 1, \dots, \dim(\phi)$ , on the bootstrap repetitions  $\hat{\phi}_{j,T:1}^*, \hat{\phi}_{j,T:2}^*, \dots, \hat{\phi}_{j,T:\tilde{B}}^*$  obtained from the estimation of the DSGE model under  $H_1$ , i.e. without imposing the CER. Then, one would expect not to reject normality on  $\hat{\theta}_{i,T:1}^*, \hat{\theta}_{i,T:2}^*, \dots, \hat{\theta}_{i,T:\tilde{B}}^*$  and on  $\hat{\phi}_{j,T:1}^*, \hat{\phi}_{j,T:2}^*, \dots, \hat{\phi}_{j,T:\tilde{B}}^*$  in DSGE models where all regularity conditions for standard inference hold. Conversely, one would expect to reject normality on  $\hat{\theta}_{i,T:1}^*, \hat{\theta}_{i,T:2}^*, \dots, \hat{\theta}_{i,T:\tilde{B}}^*$  and/or

<sup>8</sup>It would be tempting to use a Kolmogorov-Smirnov approach, i.e. to consider a functional such as  $KS_{T,\tilde{B}} := \sup_{x \in \mathbb{R}} |\tilde{B}^{1/2}V_T(x)^{-1/2}(G_{T,\tilde{B}}^*(x) - G_T^*(x))|$ . Unfortunately, in general a bootstrap statistic based on i.i.d. resampling has, for any finite sample size  $T$ , a discrete distribution with  $T^T$  atoms. Since  $G_T^*$  is discontinuous,  $KS_{\tilde{B}}$  does not have in general a pivotal limiting distribution.

on  $\hat{\phi}_{j,T:1}^*$ ,  $\hat{\phi}_{j,T:2}^*$ , ...,  $\hat{\phi}_{j,T:\tilde{B}}^*$  when the estimated DSGE model does not meet the regularity conditions for standard inference, as it happens e.g. under weak identification (or unidentification).

REMARK 4.7 The tests of (asymptotic) normality have to be computed using  $\tilde{B}$  bootstrap realizations  $\hat{\theta}_{T:1}^*$ ,  $\hat{\theta}_{T:2}^*$ , ...,  $\hat{\theta}_{T:\tilde{B}}^*$  ( $\hat{\phi}_{T:1}^*$ ,  $\hat{\phi}_{T:2}^*$ , ...,  $\hat{\phi}_{T:\tilde{B}}^*$ ) out of the (iid) sequence of  $B$  bootstrap realizations  $\hat{\theta}_{T:1}^*$ ,  $\hat{\theta}_{T:2}^*$ , ...,  $\hat{\theta}_{T:B}^*$  ( $\hat{\phi}_{T:1}^*$ ,  $\hat{\phi}_{T:2}^*$ , ...,  $\hat{\phi}_{T:B}^*$ ) employed to compute the bootstrap p-value for the QLR test of the CER. In order to control size,  $\tilde{B}$  is chosen to satisfy the condition in eq. (38) and therefore should be small compared to  $B$ . This might entail a loss of information – hence, of power – because  $B - \tilde{B}$  repetitions are discarded. For the purpose of improve power and make use of all the available information, we can partition the original  $B$  bootstrap realizations into  $N$  non-overlapping groups, each of size  $\tilde{B}$  (such that  $N\tilde{B} \leq B$ ), and run  $N$  (stochastically independent, given the original data) normality tests at the  $100\eta_0\%$  nominal significance level, one on each group. It is then reasonable to design a decision rule which summarizes the information from all  $N$  tests. Specifically, let  $p_1, \dots, p_N$  be the p-values associated with the  $N$  individual normality tests and consider the following decision rule: reject normality if  $\min_{i=1, \dots, N} \{p_i\} \leq \eta_0$ . Trivial calculations based on the stochastic independence of the  $N$  groups suggests that if each individual test is run at the  $100\eta_0\%$  nominal significance level, then the overall type I error probability is given by  $\eta = 1 - (1 - \eta_0)^N$ . Therefore one can easily select  $\eta_0$  such that  $\eta$  matches the desired overall significance level. The simulation results reported in Section 5 and in the Technical Supplement show that this indeed leads to appreciable power gains in small samples when some of the structural parameters are weakly identified.

We explore the empirical performance of the suggested bootstrap-based misspecification test for DSGE models in Sections 5 and 6 below.

#### 4.4 CONNECTIONS WITH THE LITERATURE

To our knowledge, only few papers in the literature use bootstrap methods in DSGE models. As already observed in the introduction, Chow and Moreno (2006) and Bårdsen and Fanelli (2015) exploit the potential of the bootstrap in the empirical analysis of dynamic structural models but limit their attention to the case in which all state variables are observed and do not address the consistency of the bootstrap. Bekaert and Hodrick (2001) and Fanelli and Palomba (2011) show empirically that bootstrap methods lead to substantial finite-sample size improvements in tests of the CER in the class of present value models used in financial and macroeconomic analysis, but do not prove consistency of the bootstrap. Moreover, the class of DSGE models we consider in this paper is not covered by their analysis.

Fève *et al.* (2009) suggest a minimum-distance Structural VAR (SVAR) approach for DSGE models where in a first step the model is estimated and tested by the overidentification test delivered by the minimum-distance program, and in a second-step simulated versions of the overidentification test are obtained by bootstrapping the SVAR residuals. Hence, in Fève's *et al.* (2009) bootstrap resampling involves the residuals of the auxiliary SVAR and not the residuals of the DSGE model. Similarly, Le *et al.* (2011)

combine the bootstrap with indirect inference methods for DSGE models. Compared to this literature, the distinctive feature of our approach is that we apply the bootstrap to the state space representation of the DSGE model in order to test the CER and to evaluate how ‘close’ or ‘distant’ the estimated model is from the case of strong identification.

Focusing on the more general setup of state space models, the only study which addresses the problem of bootstrap maximum likelihood estimation in state space models is the already mentioned article by Stoffer and Wall (1991); see also Berkowitz and Kilian (2000) and Stoffer and Wall (2004). Cavanaugh and Shunway (1997) exploits Stoffer and Wall’s (1991) algorithm to develop a bootstrap-based corrected variant of Akaike Information Criterion (AIC) to small sample state-space model selection. We borrow from Stoffer and Wall (1991) the design of our bootstrap algorithm, but we generalize and strengthen their results by relaxing the conditions they use to prove that the bootstrap QML estimator has the same asymptotic distribution of the QML estimator. We further prove that the bootstrap is consistent and tends to reduce the tendency of QLR tests to overreject the CER. Moreover, we show that the bootstrap can be used constructively to detect misspecification in state space models, including the case of weak identification. In this respect, Zhan (2014) is the only example where the use of the bootstrap is suggested as a diagnostic tool to detect ‘weak instruments’ in instrumental variables regressions.

Finally, there exist only few studies in the bootstrap literature where the applicability of the bootstrap is discussed in situations where not all regularity conditions for inference are assumed to hold. While the results in Moreira *et al.* (2004) suggest that the bootstrap might be valid in some weak identification cases, more recently Dovonon and Gonçalves (2014) address the bootstrap estimation of the standard test of overidentification restrictions in the generalized method of moments framework when the model is globally identified but the rank condition is not valid, a situation referred to as lack of first-order local identification.

## 5 MONTE CARLO STUDY

In this section we investigate the empirical performance of the bootstrap using An and Schorfheide’s (2007) small-scale model introduced in the Example 1 of Section 2 as data generating process (DGP). A detailed (population) identification analysis of this model may be found in e.g. Komunjer and Ng (2011). The ‘full’ vector of parameters is  $\theta := (\tau, \beta, \kappa, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_z^2, \sigma_g^2, \sigma_r^2)'$  ( $\dim(\theta) = 11$ ). Komunjer and Ng (2011) have shown that one of the three parameters of the policy rule ( $\psi_1, \psi_2, \rho_r$ ) needs to be restricted in order to achieve (population) identification, see also Qu and Tkachenko (2012). Therefore, the policy parameter  $\psi_2$  (the Central Bank’s long run response to the output gap) will be fixed at its DGP value and treated as known. Also the discount factor  $\beta$  will be fixed at its DGP value and treated as known. The DGP values of  $\theta$  are reported in Table 2 of Komunjer and Ng (2011). In this model it is difficult to characterize precisely on a priori grounds which are the candidate weakly identified parameters besides the policy parameter  $\psi_1$ . Our bootstrap misspecification approach

may play an important role in detecting weak identification in this model.

We consider two cases, denoted AS-DGP1 and AS-DGP2, respectively. In the former, it is assumed that the unknown parameters the econometrician must infer from the data are contained in  $\theta_1 := (\kappa, \rho_r, \sigma_g^2)'$  ( $\dim(\theta_1) = 3$ ), while all other elements of  $\theta$  are fixed at their DGP values. In the latter it is assumed that the unknown parameters are in  $\theta_2 := (\kappa, \psi_1, \rho_r, \sigma_e^2, \sigma_g^2, \sigma_r^2)'$  ( $\dim(\theta_2) = 6$ ), all other elements of  $\theta$  being fixed at their DGP values. In both cases, the specification of the matrices  $\Gamma_0$ ,  $\Gamma_f$ ,  $\Gamma_b$ ,  $\Pi$  and  $\Psi$  in the structural representation of the DSGE model is given in Example 1 and the implied state space representation without the CER imposed involves the vector  $\phi$ , where  $\dim(\phi) = 13$ . AS-DGP1 represents our benchmark for the strong identification case, while we expect to find identification issues in model AS-DGP2, where the policy rule parameters  $\psi_1$ ,  $\rho_r$  and  $\sigma_r^2$ , according to an extensive literature, are difficult to identify from the data; see e.g. Mavroeidis (2010) and Castelnuovo and Fanelli (2015).

We generate samples of length  $T = 100$  from the minimal ABCD form associated with the DSGE model  $M = 500$  times, assuming that the vector of fundamental shocks  $\omega_t := (\varepsilon_{e,t}, \varepsilon_{g,t}, \varepsilon_{r,t})'$  ( $n_\omega = 3$ ) is Gaussian with diagonal covariance matrix. The initial conditions  $Z_{m,0}$  are fixed to zero. For each replication, a sample of  $T + 200$  observations is actually generated and the first 200 observations are then discarded. We then analyze (i) the empirical size of the QLR test for the CER, (ii) the bootstrap standard errors and the bootstrap-adjusted rejection frequency of the QLR test for the CER obtained with the algorithm discussed in Section 4.1 and (iii) the empirical performance of our bootstrap misspecification approach discussed in Section 4.3 for different choices of  $\tilde{B} \leq B$ . In (iii) we complement the analysis by applying the procedure summarized in Remark 4.7.

Since the data are simulated assuming a Gaussian distribution for the shocks, for the rest of this section we use ‘ML’ and ‘LR’ in place of ‘QML’ and ‘QLR’, respectively.

## 5.1 AS-DGP1

Results relative to AS-DGP1 are summarized in Table 1 ( $T = 100$ ) (see the Technical Supplement for the case  $T = 250$ ), which summarizes the situation one would expect to find in strongly identified DSGE models. Averaging over the MC replications, the ML estimates of the structural parameters and Hessian-based standard errors are substantially similar to the bootstrap ML estimates and bootstrap standard errors, as expected in the case where the regularity conditions hold. When asymptotic critical values taken from the  $\chi_k^2$  distribution,  $k = \dim(\phi) - \dim(\theta_1) = 10$ , are employed, the rejection frequency (empirical size) of the LR test for the CER is 8% as opposed to the 5% nominal significance level. The bootstrap is more conservative, with an associated rejection frequency around 2%. Therefore, the bootstrap seems to counterbalance the tendency of the standard LR test to overreject the CER, although it leads to a slightly conservative test.

Focusing on the bootstrap misspecification approach, in the bottom panel of Table 1 we report rejection frequencies obtained using Jarque and Bera’s (1987) univariate test of normality of the  $\tilde{B}$  i.i.d. realizations of the bootstrap estimators;<sup>9</sup> we also report the

<sup>9</sup>We also applied Shapiro and Wilk’s (1965) normality test obtaining substantially similar results.

rejection frequencies associated with the procedure summarized in Remark 4.7. Recall from Section 4.1 that according to the theory, the number of bootstrap repetitions  $\tilde{B}$  ( $\tilde{B} \leq B$ ) used for testing (asymptotic) normality of the bootstrap estimator, should be small compared to  $T$ , see eq. (38). Hence, in order to evaluate how the condition in eq. (38) impacts on the finite sample performance of the test, we consider different choices of  $\tilde{B} \leq B$ . Normality tests are then computed at the 5% nominal significance level. We observe that for values of  $\tilde{B}$  sufficiently small relative to  $T$ , the rejection frequencies of the tests fluctuate around the nominal size. On the other hand, as  $\tilde{B}$  increases relative to  $T$ , the rejection frequency of the test tends to exceed the nominal significance level, showing that choosing  $\tilde{B}$  too large can lead to spurious rejections of the hypothesis of asymptotic normality of the bootstrap estimators. In summary, these results confirm that, when  $\tilde{B}$  is not unreasonable large relative to  $T$ , our misspecification test based on normality tests delivers a good empirical size coverage in strongly identified DSGE models.

## 5.2 AS-DGP2

Results relative to AS-DGP2 are summarized in Table 2 (see the Technical Supplement for the case  $T = 250$ ). As expected, compared to the previous case Table 2 depicts a more involved situation. After averaging over the MC replications, the ML estimates of the structural parameters and their bootstrap counterparts have similar magnitudes. However, let us consider the distance between the asymptotic and the bootstrap standard errors, which can be regarded as a ‘descriptive’ indicator of weak identification. It can be clearly noticed that for the policy rule parameters  $\psi_1$  and  $\rho_r$  there are non-negligible discrepancies between the Hessian-based and the bootstrap standard errors. Such discrepancies do not vanish when the sample size is increased to  $T = 250$  (see in particular Table TS.12 in the Technical Supplement). Table 2 also shows that using the asymptotic critical values taken from the  $\chi_k^2$  distribution,  $k = \dim(\phi) - \dim(\theta_2) = 7$ , the rejection frequency of the LR test for the CER is 7.8% as opposed to the 5% nominal significance level, while the bootstrap-adjusted rejection frequency is 3.2%.

As concerns our bootstrap misspecification approach based on normality tests, from the bottom panel of Table 2 we notice two facts. First, we observe possible identification issues for the parameters  $\kappa$  (slope of the Phillips curve) and  $\psi_1$  (long run policy response to inflation) because the rejection frequencies associated with the normality tests of  $\hat{\kappa}_T^*$  and  $\hat{\psi}_{1,T}^*$  tend to exceed the 5% nominal significance markedly relative to what happens to the other parameters. Second, the rejection frequencies associated with  $\hat{\kappa}_T^*$  and  $\hat{\psi}_{1,T}^*$  increase significantly when the procedure summarized in Remark 4.7 is employed. Based on the evidences in Table 2, we can conclude that estimation of  $\theta_2 := (\kappa, \psi_1, \rho_r, \sigma_e^2, \sigma_g^2, \sigma_r^2)'$  in model AS-DGP2 is characterized by weak identification issues which can be mainly ascribed to the policy parameter  $\psi_1$ . However, according to our misspecification tests, the deviations from standard regularity conditions do not seem dramatic. Interestingly, despite the presence of weakly identified parameters, we report the same phenomenon noticed for the strongly identified AS-DGP1 model: the bootstrap counterbalances the tendency of the LR test to overreject the CER when asymptotic critical values are used. The additional Monte Carlo experiments in the

Technical Supplement confirm this evidence.

## 6 EMPIRICAL ILLUSTRATION

In this section we apply our bootstrap approach to estimate and evaluate An and Schorfheide’s (2007) DSGE model of Example 1 on U.S. data. We employ quarterly observations relative to the ‘Great Moderation’ sample 1984Q2-2008Q3,  $T=98$ . The starting date, 1984Q2, is justified by McConnell and Pérez-Quirós (2000), who find a break in the variance of the U.S. output growth in 1984Q1. The ending date is instead motivated by the fact that, with data after 2008Q3, it would be hard to identify a ‘conventional’ monetary policy shock with our structural model during the well known zero lower bound (ZLB) episodes. The three observable variables in  $y_t := (\tilde{x}_t, \pi_t, R_t)'$  are measured as follows. The output gap,  $\tilde{x}_t$ , is computed as percent log-deviation of the real GDP from the potential output, as estimated by the Congressional Budget Office. The inflation rate,  $\pi_t$ , is the quarterly growth rate of the GDP deflator. For the short-term nominal interest rate,  $R_t$ , we consider the effective Federal funds rate expressed in quarterly terms (averages of monthly values). The source of the data is the Federal Reserve Bank of St. Louis’ web site.

As in the Monte Carlo study, we consider two estimable versions of the model. In the former, denoted AS-M1, the vector of unknown parameters is  $\theta_1 := (\kappa, \rho_r, \sigma_g^2)'$  ( $\dim(\theta_1)=3$ ) and all other structural parameters of system (3)-(8) are calibrated at the values reported in the upper panel of Table 1 (hence they are assumed to be known by the econometrician). In the latter, denoted AS-M2, the vector of unknown parameters is  $\theta_2 := (\kappa, \psi_1, \rho_r, \sigma_z^2, \sigma_g^2, \sigma_r^2)'$  ( $\dim(\theta_2)=6$ ), with all other parameters characterizing system (3)-(8) calibrated at the values reported in the upper panel of Table 1 (again, these are assumed to be known to the econometrician). The main difference between model AS-M1 and model AS-M2 is that the latter estimates the Central Bank’s long run response to inflation  $\psi_1$  and the variance of the three fundamental shocks  $\sigma_z^2, \sigma_g^2, \sigma_r^2$  from the data, while AS-M1 treats  $\psi_1$  and  $\sigma_e^2$  and  $\sigma_r^2$  as fixed. Thus, AS-M1 is a restricted version of AS-M2. It is worth noting that the two models share the same state space representation (and likelihood) under  $H_1$ , i.e. when the CER are not imposed.

We run some preliminary tests of no autocorrelation and heteroskedasticity in the innovation residuals obtained from the (Gaussian) Kalman filter estimation of the DSGE state space representation, both under  $H_0$  and  $H_1$ . We find some residual heteroskedasticity, especially in the measurement equation associated to the short-term interest rate. This result suggests that also the wild-bootstrap version of our QLR test could be considered, see Remark 4.3.

Estimation is carried out by combining the Kalman filter with the ‘CMA-ES’ likelihood maximization algorithm (Andreasen, 2010). For both models, the bootstrap standard errors associated with the QML estimates of the structural parameters and the bootstrap p-values associated with the LR tests for the CER are computed using  $B = 999$  replications, following the algorithm summarized in Section 4.1. Our misspecification tests are instead based on  $\tilde{B} = 20$  bootstrap realizations of the parameter estimators and are constructed as described in Section 4.3 using Jarque and Bera’s



(1987) normality test at the 5% nominal significance level. The choice  $\tilde{B} = 20$  is motivated by the requirement in eq. (38). We complement the p-values associated with the normality tests with the results from the procedure described in Remark 4.7.

Results are summarized in Table 3. The upper panel of Table 3 refers to model AS-M1. We notice that in this case, the bootstrap parameter estimator  $\hat{\theta}_{1,T}^*$  (fourth column), calculated as  $(1/B) \sum_{b=1}^B \hat{\theta}_{1,T:b}^*$ , is close to the QML estimate computed on the original sample  $\hat{\theta}_{1,T}$  (second column). The largest difference is seen for the policy rule inertia parameter  $\rho_r$ . Similarly, the bootstrap standard errors (fifth column) are similar to the Hessian-based standard errors (third column) and, again, the largest difference characterizes the inertia parameter  $\rho_r$ . Thus, a descriptive indicator of weak identification – based on the difference between analytic and bootstrap-based estimated standard errors – does not signal relevant weak identification issues. Using asymptotic critical values (taken from the  $\chi_k^2$  distribution, where  $k = \dim(\phi) - \dim(\theta_1) = 13 - 3 = 10$ ), the QLR test rejects the CER at the 5% nominal significance level. Instead, the (iid) bootstrap p-value associated with the QLR test is equal to 0.053, while the wild-bootstrap p-value is 0.147.

The (asymptotic) normality of the bootstrap estimators of the structural parameters is not rejected for any parameter at the 5% nominal significance level.<sup>10</sup> We have enough ingredients to conclude that model AS-M1 satisfies the condition for standard asymptotic inference and, notably, is not rejected at the 5% nominal significance level when we consider the iid and wild bootstrap version of the QLR test for the CER, while it is rejected using first-order asymptotics. Interestingly, our Technical Supplement shows that model AS-M1 is rejected using Guerron-Quintana’s *et al.* (2013) identification-robust method.

The lower panel of Table 3 refers to model AS-M2. We observe that in this model the magnitudes of the QML estimates of  $\kappa$  (slope of the Phillips curve) and of  $\sigma_g^2$  (variance of the shock in the Phillips curve) are substantially different from the magnitudes estimated in model AS-M1. In particular, treating  $\psi_1, \sigma_r^2$  and  $\sigma_z^2$  as estimable parameters of the DSGE model renders the magnitude of the QML estimate of the slope parameter  $\hat{\kappa}_T$  very large. Moreover, the QML estimate of  $\psi_1$  lies on the boundary of the optimization parameter space; see the notes to Table 3 for details and condition (A.v\*) in the Appendix. Compared to model AS-M1, the discrepancy between  $\hat{\theta}_{2,T}^*$  (fourth column) and  $\hat{\theta}_{2,T}$  (second column) is larger and can be essentially ascribed to the estimates of the parameters  $\rho_r$  and  $\sigma_g^2$ . Considering our ‘descriptive’ indicator of weak identification, the bootstrap standard errors (fifth column) diverge substantially from the Hessian-based standard errors (third column). Using asymptotic critical values (taken from the  $\chi_k^2$  distribution, where  $k = \dim(\phi) - \dim(\theta_2) = 13 - 6 = 7$ ), the QLR test rejects the CER at the 5% nominal significance level. Also the iid bootstrap p-value rejects the CER, while the wild-bootstrap p-value is 0.112. The difference between the

---

<sup>10</sup>In this empirical illustration we do not consider normality tests on the elements of  $\hat{\phi}_T^*$ , i.e. on the bootstrap ML estimators of the parameters of the DSGE model under  $H_1$ . One fact to be noticed is that the absolute value of the largest eigenvalue of the estimated matrix  $A(\hat{\phi}_T)$  is 0.947, which suggests that the system features persistent variables. In this specific case, the largest eigenvalue of  $A(\hat{\phi}_T)$  coincides with one of the elements of  $\hat{\phi}_T$ , say  $\hat{\phi}_{1,T}$ , hence it is reasonable to expect the rejection of normality for  $\hat{\phi}_{1,T}^*$ . Results are available upon request.

iid and wild bootstrap can be explained by the heteroskedasticity found in the innovations residuals. Asymptotic normality of the realizations of the bootstrap structural parameter estimators is sharply rejected for  $\hat{\kappa}_T^*$ ,  $\hat{\psi}_{1,T}^*$  and  $\hat{\sigma}_{z,T}^{2*}$  at the 5% nominal significance level.<sup>11</sup> We have clear evidence that, compared to model AS-M1, the regularity conditions for standard inference are not respected in model AS-M2, which seems to be affected by identification issues.

## 7 CONCLUSIONS

In this paper we have explored the potential of bootstrap methods in the empirical evaluation of DSGE models. We have shown that with our bootstrap approach, finite sample inference on the CER implied by DSGE models is considerably more accurate and the risk of false rejections can be reduced also when part of the structural parameters are weakly identified. We have further proposed a novel bootstrap misspecification approach which can be used to analyze the strength of identification of the estimated model when all other possible causes of misspecification can be controlled for. This test seems to be promising, based on the evidence produced in the paper, and straightforward to apply. Our results contrast the common wisdom that the bootstrap is not useful in DSGE models.

## A APPENDIX: CONSISTENCY OF THE BOOTSTRAP IN STRONGLY IDENTIFIED DSGE MODELS

In this Appendix, we provide heuristic proofs of Proposition 1 and Proposition 2. Our proofs include two short preliminary sections where we fix the notation necessary to understand the bootstrap setting and we define the objective functions.

### A.1 NOTATION

With  $P(\cdot)$  denoting the true probability distribution of the data, we use  $E(\cdot)$  and  $\text{Var}(\cdot)$ , respectively, to denote expectations and variance computed under  $P$ . We use  $P^*(\cdot)$  to denote the probability measure induced by the bootstrap, conditional on the original sample. Expectation and variance computed under  $P^*$  are denoted by  $E^*(\cdot)$  and  $\text{Var}^*(\cdot)$ , respectively. In general, with respect to the original probability measure,  $P^*(\cdot)$ ,  $E^*(\cdot)$  and  $\text{Var}^*(\cdot)$  are stochastic quantities. Define, for any  $\delta > 0$ ,  $p_T := P(\|\hat{\theta}_T - \theta_0\| > \delta)$  and  $p_T^* := P^*(\|\hat{\theta}_T^* - \hat{\theta}_T\| > \delta)$ , where  $\hat{\theta}_T^*$  is the bootstrap analog of the QML estimator  $\hat{\theta}_T$ , and  $\|\cdot\|$  is the Euclidean norm. While with the conventional notation ‘ $\hat{\theta}_T - \theta_0 \xrightarrow{P} 0$ ’ we mean that the (deterministic) sequence  $\{p_T\}$  converges to zero ( $p_T \rightarrow 0$ ) as  $T \rightarrow \infty$ , with the notation ‘ $\hat{\theta}_T^* - \hat{\theta}_T \xrightarrow{P^*} 0$ ’, which reads ‘ $\hat{\theta}_T^* - \hat{\theta}_T$  converges in  $P^*$  to 0, in probability’, we mean that the (stochastic) sequence  $\{p_T^*\}$  converges in probability to zero ( $p_T^* \xrightarrow{P^*} 0$ ).

---

<sup>11</sup>The rejection of the (asymptotic) normality of  $\hat{\psi}_{1,T}^*$  is not surprising. Indeed  $\hat{\psi}_{1,T}^*$  lies on the boundary of the parameter space, violating one of the conditions (see point A.v\*, Appendix A) at the basis of the Gaussian asymptotic approximation.

Likewise, for any  $\delta > 0$ , define  $s_T := P(\sup_{\theta \in \mathcal{P}^D} |\hat{Q}_T(\theta) - Q_0(\theta)| > \delta)$ , where  $\hat{Q}_T(\theta)$  and  $Q_0(\theta)$  are the criterion functions defined below, and let  $s_T^* := P^*(\sup_{\theta \in \mathcal{P}^D} |\hat{Q}_T^*(\theta) - Q_0(\theta)| > \delta)$ , where  $\hat{Q}_T^*(\theta)$  is the bootstrap analog of  $\hat{Q}_T(\theta)$  given the original sample, see below. While ‘ $\hat{Q}_T(\theta)$  converges uniformly in probability to  $Q_0(\theta)$ ’ means that the (deterministic) sequence  $\{s_T\}$  converges to zero ( $s_T \rightarrow 0$ ), with ‘ $\hat{Q}_T^*(\theta)$  converges uniformly in  $P^*$  to  $Q_0(\theta)$ , in probability’ we mean that the stochastic sequence  $\{s_T^*\}$  converges to zero in probability ( $s_T^* \xrightarrow{p} 0$ ). Alternative useful notations are  $\sup_{\theta \in \mathcal{P}^D} |\hat{Q}_T(\theta) - Q_0(\theta)| \xrightarrow{p} 0$  (uniform convergence in probability) and  $\sup_{\theta \in \mathcal{P}^D} |\hat{Q}_T^*(\theta) - Q_0(\theta)| \xrightarrow{p^*} 0$  (uniform convergence in  $P^*$ , in probability), respectively.

Finally, consider a scalar stochastic sequence  $\{W_T\}$  and a random variable  $W$ , with associated CDFs denoted by  $G_{W,T}(x) := P(W_T \leq x)$  and  $G_W(x) := P(W \leq x)$ . The conventional notation ‘ $W_T \xrightarrow{d} W$ ’ means that  $G_{W,T}(x) \rightarrow G_W(x)$  for each  $x$  at which  $G_W(x)$  is continuous. Let the sequence  $\{W_T^*\}$  be the bootstrap analog of  $\{W_T\}$ , with associated CDFs  $G_{W^*,T}^*(x) := P^*(W_T^* \leq x)$ . We say that  $W_T^*$  ‘converges in conditional distribution to  $W$ , in probability’, denoted by ‘ $W_T^* \xrightarrow{d^*} W$ ’, if  $G_{W^*,T}^*(x) \xrightarrow{p} G_W(x)$  for each  $x$  at which  $G_W(x)$  is continuous. Notice that if  $G_W(\cdot)$  is continuous, then the latter convergence also implies that  $\sup_{x \in \mathbb{R}} |G_{W^*,T}^*(x) - G_W(x)| \xrightarrow{p} 0$ . These definitions can be extended to the multivariate framework in the conventional way.

## A.2 OBJECTIVE FUNCTIONS: DEFINITIONS

Consider  $\ell_{\circ,T}(\theta) = \ell_T(\phi_\theta)$  the essential part of the log-likelihood function associated with the DSGE model, see Section 3.1. We define the normalized log-likelihood:

$$\hat{Q}_T(\theta) := \frac{1}{T} \ell_{\circ,T}(\theta) = \frac{1}{T} \sum_{t=1}^T \ell_t(\theta) \quad (\text{A.1})$$

where  $\ell_t(\theta) := l(y_t | \mathcal{F}_{t-1}^y; \theta) = -\{\log \det(\Sigma_{\epsilon^0,t}(\theta)) + e_t^0(\theta)' e_t^0(\theta)\}$ ,  $e_t^0(\theta) := \Sigma_{\epsilon^0,t}^{-1/2}(\theta) \epsilon_t^0(\theta)$  (see Technical Supplement). The QML estimator  $\hat{\theta}_T$  solves  $\hat{\theta}_T := \arg \max_{\theta \in \mathcal{P}^D} \hat{Q}_T(\theta)$ . By construction,  $\hat{Q}_T(\theta)$  is twice continuously differentiable in a neighborhood  $\mathcal{N}_{\theta_0}$  of  $\theta_0$ . Associated with  $\hat{Q}_T(\theta)$  we have

$$Q_T(\theta) := E(\hat{Q}_T(\theta)) = E\left(\frac{1}{T} \sum_{t=1}^T \ell_t(\theta)\right) = \frac{1}{T} \sum_{t=1}^T E(\ell_t(\theta)) = E(\ell_t(\theta)) =: Q_0(\theta) \quad (\text{A.2})$$

where the last equality holds under Assumption 1 (which implies that  $\{y_t\}$  is weakly stationary).

In the bootstrap world, the analog of  $\hat{Q}_T(\theta)$  is given by

$$\hat{Q}_T^*(\theta) := \frac{1}{T} \ell_{\circ,T}^*(\theta) = \frac{1}{T} \sum_{t=1}^T \ell_t^*(\theta) \quad (\text{A.3})$$

where  $\ell_t^*(\theta) := l^*(y_t^* | \mathcal{F}_{t-1}^{y^*}; \theta) = -\{\log \det(\Sigma_{\epsilon^0,t}(\theta)) + e_t^{0,*}(\theta)' e_t^{0,*}(\theta)\}$ ,  $e_t^{0,*}(\theta)$ ,  $t = 1, \dots, T$ , being the bootstrap Kalman filter innovations. Obviously,  $\hat{\theta}_T^* := \arg \max_{\theta \in \mathcal{P}^D} \hat{Q}_T^*(\theta)$ .

Associated with  $\hat{Q}_T^*(\theta)$  we have its (conditional) expectation

$$Q_T^*(\theta) := E^*(\hat{Q}_T^*(\theta)) = E^* \left( \frac{1}{T} \sum_{t=1}^T \ell_t^*(\theta) \right) = \frac{1}{T} \sum_{t=1}^T E^*(\ell_t^*(\theta)) = E^*(\ell_t^*(\theta)) =: Q_0^*(\theta). \quad (\text{A.4})$$

### A.3 CONSISTENCY OF THE BOOTSTRAP QML ESTIMATOR

Throughout this section, we denote with (TS-A.j),  $j = \text{i, ii, iii, iv}$ , the regularity conditions summarized in the Technical Supplement. To prove the consistency in eq. (32), we proceed by showing that, in probability as  $T \rightarrow \infty$ , we have bootstrap analogs of (TS-A.i)-(TS-A.iv) that are valid for the criterion function  $\hat{Q}_T^*(\theta)$  evaluated at  $\hat{\theta}_T$ .

We need the Lemma that follows.

LEMMA A.1 (STOFFER AND WALL, 1991) *Under the regularity conditions (TS-A.i)-(TS-A.iv),*

$$Q_T^*(\theta) = \hat{Q}_T(\theta) \quad \text{for all } \theta \in \mathcal{P}^D$$

PROOF: See Stoffer and Wall (1991). ■

By Lemma 1,  $\hat{\theta}_T$  also equals  $\arg \max_{\theta \in \mathcal{P}^D} Q_0^*(\theta)$ ; hence, in the bootstrap world  $\hat{\theta}_T$  is the unique maximizer of  $Q_0^*(\theta)$  in a neighborhood  $\mathcal{N}_{\hat{\theta}_T}$  of  $\hat{\theta}_T$ . This condition is the bootstrap analog of (TS-A.ii), when (TS-A.ii) is restricted in a neighborhood  $\mathcal{N}_{\theta^\dagger}$  of  $\theta^\dagger$ . Also by Lemma 1,  $Q_0^*(\theta)$  is continuous in  $\theta$  since so is  $\hat{Q}_T(\theta)$ . This is the bootstrap analog of (TS-A.iii). If in the presence of stationarity (and ergodicity)  $\{\ell_t(\theta)\}$  satisfies a UWLLN and also the bootstrap analog  $\{\ell_t^*(\theta)\}$  satisfies a UWLLN in  $P^*$  in probability, i.e.

$$P^* \left( \sup_{\theta \in \mathcal{P}^D} \left| \hat{Q}_T^*(\theta) - Q_0^*(\theta) \right| \right) \xrightarrow{p} 0 \text{ as } T \rightarrow \infty,$$

then, by Lemma 1,

$$P^* \left( \sup_{\theta \in \mathcal{P}^D} \left| \hat{Q}_T^*(\theta) - \hat{Q}_T(\theta) \right| \right) \xrightarrow{p} 0 \text{ as } T \rightarrow \infty. \quad (\text{A.5})$$

Eq. (A.5) is the bootstrap analog of (TS-A.iv).

Summing up, we have the following bootstrap analogs of the regularity conditions (TS-A.i)-(TS-A.iv) reported in the Technical Supplement:

(A.i\*)  $\mathcal{P}^D$  is compact;

(A.ii\*)  $\hat{\theta}_T$  is the unique maximizer of  $Q_0^*(\theta)$  in a neighborhood  $\mathcal{N}_{\hat{\theta}_T}$  of  $\hat{\theta}_T$ ;

(A.iii\*)  $Q_0^*(\theta)$  is continuous in  $\theta$ ;

(A.iv\*)  $\hat{Q}_T^*(\theta) - \hat{Q}_T(\theta)$  converges uniformly in  $P^*$  in probability to 0, i.e. eq. (A.5)

holds.

It turns out that

$$\hat{\theta}_T^* - \hat{\theta}_T \xrightarrow{p^*} 0 \quad (\text{A.6})$$

where  $\hat{\theta}_T^* := \arg \max_{\theta \in \mathcal{N}_{\hat{\theta}_T} \cap \mathcal{P}^D} \hat{Q}_T^*(\theta)$  and  $\hat{\theta}_T := \arg \max_{\theta \in \mathcal{N}_{\hat{\theta}_T} \cap \mathcal{P}^D} \hat{Q}_T(\theta)$ . This proves the convergence in eq. (32) of part (i) of Proposition 1.

#### A.4 ASYMPTOTIC NORMALITY OF THE BOOTSTRAP QML ESTIMATOR

Throughout this section, we denote with (TS-A.*j*), *j* = v, vi, vii, viii, the regularity conditions summarized in the Technical Supplement. To prove the asymptotic normality of  $Z_T := T^{1/2}(\hat{\theta}_T^* - \hat{\theta}_T)$ , we consider a mean value expansion of  $\nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T^*) = 0_{\dim(\theta) \times 1}$  around the QML estimator  $\hat{\theta}_T$ , yielding

$$\nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T^*) = 0_{\dim(\theta) \times 1} = \nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T) + \nabla_{\theta\theta}^2 \hat{Q}_T^*(\bar{\theta}_T^*)(\hat{\theta}_T^* - \hat{\theta}_T)$$

where  $\bar{\theta}_T^*$  is on the segment connecting the points  $\hat{\theta}_T^*$  and  $\hat{\theta}_T$  of  $\mathcal{P}^D$ . Rearranging terms,

$$\begin{aligned} T^{1/2}(\hat{\theta}_T^* - \hat{\theta}_T) &= - \left\{ \nabla_{\theta\theta}^2 \hat{Q}_T^*(\bar{\theta}_T^*) \right\}^{-1} T^{1/2} \nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T) \\ &= - \left\{ T^{-1} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t^*(\bar{\theta}_T^*) \right\}^{-1} T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T). \end{aligned} \quad (\text{A.7})$$

To prove that the term on the right in eq. (A.7) converges to the Gaussian distribution, we need a bootstrap analog of the CLT in (TS-A.vi) and we have to prove that as  $T \rightarrow \infty$ , the implied long run covariance matrix is indistinguishable from  $\mathcal{B}(\hat{\theta}_T) := \text{Var}(T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t(\hat{\theta}_T))$ . Moreover, in order to derive the limiting behaviour of the left term in eq. (A.7), we need a bootstrap analog of the uniform convergence result of the Hessian matrix in (TS-A.vii), and a bootstrap analog of the nonsingularity condition (TS-A.viii).

As concerns the bootstrap analog of (TS-A.vi), if  $T^{1/2} \nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T) = T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T)$  satisfies a CLT, we then have, with  $\mathcal{B}^*(\hat{\theta}_T) := \text{Var}^*(T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T))$ :

$$T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T) \xrightarrow{d^*} N(0_{\dim(\theta) \times 1}, \mathcal{B}^*) \quad (\text{A.8})$$

where  $\mathcal{B}^* := \text{plim}_{T \rightarrow \infty} \mathcal{B}^*(\hat{\theta}_T)$ . We can specialize this result using the following Lemma, which is a reformulation of Lemma 2 in Stoffer and Wall (1991).

LEMMA A.2 (STOFFER AND WALL, 1991) *Consider the regularity conditions (TS-A.i)-(TS-A.viii). With  $\mathcal{B}^*(\hat{\theta}_T)$  as defined above and*

$$\mathcal{B}(\hat{\theta}_T) := \text{Var} \left( T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t(\hat{\theta}_T) \right),$$

it holds that

$$\mathcal{B}^*(\hat{\theta}_T) - \mathcal{B}(\hat{\theta}_T) \xrightarrow{p^*} 0.$$

As concerns the bootstrap analog of the uniform convergence result of the Hessian matrix in (TS-A.vii), if similarly to  $\{\nabla_{\theta\theta}^2 \ell_t(\theta)\}$  also – for  $T$  large enough – the (stationary and ergodic) process  $\{\nabla_{\theta\theta}^2 \ell_t^*(\theta)\}$  satisfies a UWLLN, then

$$\sup_{\theta \in \mathcal{N}_{\hat{\theta}_T}} \left\| \nabla_{\theta\theta}^2 \hat{Q}_T^*(\theta) - E^*(\nabla_{\theta\theta}^2 \hat{Q}_T^*(\theta)) \right\| \xrightarrow{p^*} 0. \quad (\text{A.9})$$

By Lemma A.1 we have that

$$\begin{aligned} E^*(\nabla_{\theta\theta}^2 \hat{Q}_T^*(\theta)) &= E^* \left( \frac{1}{T} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t^*(\theta) \right) = \frac{1}{T} \sum_{t=1}^T E^* (\nabla_{\theta\theta}^2 \ell_t^*(\theta)) \\ &= E^* (\nabla_{\theta\theta}^2 \ell_t^*(\theta)) = \frac{1}{T} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t(\theta) =: \mathcal{A}_T(\theta) \end{aligned}$$

for all  $\theta \in \mathcal{P}^D$ . Therefore, since as  $T \rightarrow \infty$  we have that  $\hat{\theta}_T - \theta^\dagger \xrightarrow{p^*} 0$ , it also holds that, uniformly in a neighborhood of  $\hat{\theta}_T$ ,  $\mathcal{A}_T(\theta) - \mathcal{A}(\theta) \rightarrow_p 0$ , where  $\mathcal{A}(\theta)$ , defined in the Technical Supplement, is continuous at  $\hat{\theta}_T$ . This result, along with (A.9), also implies that

$$\frac{1}{T} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t^*(\hat{\theta}_T) - \mathcal{A}(\hat{\theta}_T) \xrightarrow{p^*} 0$$

where  $\mathcal{A}(\hat{\theta}_T)$  can be estimated consistently by  $\mathcal{A}_T(\hat{\theta}_T) := -T^{-1} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t(\hat{\theta}_T)$ .

Summing up, we have the following bootstrap analogs of the regularity conditions (TS-A.v)-(TS-A.viii):

- (A.v\*)  $\hat{\theta}_T$  is in  $\text{int}(\mathcal{P}^D)$  with probability tending to one as  $T \rightarrow \infty$ ;
- (A.vi\*)  $T^{1/2} \nabla_{\theta} \hat{Q}_T^*(\hat{\theta}_T) = T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T)$  satisfies the CLT:

$$T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T) \xrightarrow{d^*} N(0_{\dim(\theta) \times 1}, \mathcal{B}^*) \quad , \quad \mathcal{B}^* := \mathcal{B}^*(\hat{\theta}_T), \quad (\text{A.10})$$

where  $\mathcal{B}^*(\hat{\theta}_T) := \text{Var}^* \left( T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t^*(\hat{\theta}_T) \right)$ ;

- (A.vii\*)  $\sup_{\theta \in \mathcal{N}_{\hat{\theta}_T}} \left\| \nabla_{\theta\theta}^2 \hat{Q}_T^*(\theta) - E^*(\nabla_{\theta\theta}^2 \hat{Q}_T^*(\theta)) \right\| \xrightarrow{p^*} 0$ ;
- (A.viii\*)  $\mathcal{A}(\hat{\theta}_T)$  is nonsingular, almost surely.

Going back to the mean value expansion in eq. (A.7), we have that the first term on the left is close (in  $P^*$ -in probability) to  $\mathcal{A}(\hat{\theta}_T)$  because of (A.vii\*)-(A.viii\*) and the result in eq. (A.6). The right-hand side term satisfies the convergence in eq. (A.10) by (A.vi\*). It follows that

$$[\mathcal{A}(\hat{\theta}_T)^{-1} \mathcal{B}(\hat{\theta}_T) \mathcal{A}(\hat{\theta}_T)^{-1}]^{-1/2} \{T^{1/2}(\hat{\theta}_T^* - \hat{\theta}_T)\} \xrightarrow{d^*} N(0_{\dim(\theta) \times 1}, I_{\dim(\theta)}). \quad (\text{A.11})$$

This proves the convergence in eq.s (33)-(34) of part (i) of Proposition 1. Observe that the quantity  $\mathcal{A}(\hat{\theta}_T)^{-1} \mathcal{B}(\hat{\theta}_T) \mathcal{A}(\hat{\theta}_T)^{-1}$  in eq. (A.11) converges in probability to  $V_{\theta^\dagger}$ .

The asymptotic normality in eq. (A.11) implies that the statistic  $K_T^* := T^{1/2} Z_T^* := T^{1/2} \tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{-1/2} (\hat{\theta}_T^* - \hat{\theta}_T)$  which has CDF  $G_T^*(x) := P^*(K_{1,T}^* \leq x_1, \dots, K_{\dim(\theta), T}^* \leq x_{\dim(\theta)})$  for finite  $T$ , converges asymptotically to  $G_\infty^*(x) = \Phi(x)$ . From Polya's theorem we obtain the consistency result in eq. (35) of part (ii) of Proposition 1. ■

## REFERENCES

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

- Anderson, B.D.O. and Moore, J. (1979), *Optimal filtering*, Englewood Cliffs, NJ: Prentice-Hall.
- Andreasen, M. M. (2010), How to maximize the likelihood function for a DSGE model, *Computational Economics* 35, 127-154.
- Andrews, D.W.K., Cheng, X. (2012), Estimation and inference with weak, semi-strong, and strong identification, *Econometrica* 80, 2153-2211.
- Andrews, I., Mikusheva, A. (2015), Maximum likelihood inference in weakly identified dynamic stochastic general equilibrium models, *Quantitative Economics* 6, 123-152.
- Angelini, G., Cavaliere, G. and Fanelli, L. (2016), Technical supplement to ‘Bootstrapping DSGE models’, available at [www.rimini.unibo.it/fanelli/TS\\_bootstrapping\\_DSGE.pdf](http://www.rimini.unibo.it/fanelli/TS_bootstrapping_DSGE.pdf).
- Angelini, G., Fanelli, L. (2015), Misspecification and expectations correction in New-Keynesian DSGE models, *Oxford Bulletin of Economics and Statistics*, forthcoming.
- Bårdsen, G., Fanelli, L. (2015), Frequentist evaluation of small DSGE models, *Journal of Business and Economic Statistics* 33, 307-322.
- Bekaert G., Hodrick, R. (2001), Expectations hypotheses tests, *Journal of Finance* 56, 1357-1394.
- Berkowitz, J., Kilian, L. (2000), Recent developments in bootstrapping time series, *Econometric Reviews* 19, 1-48.
- Boswijk, P., Cavaliere, G., Rahbek, A. and Taylor, R. A.M. (2015), Inference on Co-integration Parameters in Heteroskedastic Vector Autoregressions, *Journal of Econometrics*, forthcoming.
- Canova F., Sala, L. (2009), Back to square one: identification issues in DSGE models, *Journal of Monetary Economics* 56, 431-449.
- Castelnuovo, E., Fanelli, L. (2015), Monetary policy indeterminacy and identification failures in the U.S.: Results from a robust test, *Journal of Applied Econometrics* 30, 857-1010.
- Cavaliere, G., Rahbek, A., and Taylor R. A. M. (2012), Bootstrap Determination of the Co-Integration Rank in Vector Autoregressive Models, *Econometrica* 80, 1721-1740.
- Cavaliere, G., Rahbek, A. and Nielsen, H. B. (2015a), Bootstrap testing of hypotheses on co-integration relations in VAR models, *Econometrica* 83, 813–831.
- Cavaliere, G., Nielsen, H.B. and Rahbek, A. (2015b), On the consistency of bootstrap testing for a parameter on the boundary of the parameter space, mimeo.

- Cavanaugh, J.E. and Shumway, R.H. (1997), A bootstrap variant of AIC for state-space model selection, *Statistica Sinica* 7, 473-496.
- Cho, S., Moreno, A. (2006), A small-Sample study of the New-Keynesian macro model, *Journal of Money Credit and Banking* 38, 1462-1482.
- Davidson, R., MacKinnon, J. G. (1999), The size distortion of bootstrap tests, *Econometric Theory* 15, 361-376.
- Diebold, F. X., Ohanian, L. E., Berkowitz, J. (1998), Dynamic equilibrium economies: a framework for comparing models and data, *Review of Economic Studies* 65, 433-452.
- Dovonon, P. and Gonçalves, S. (2014), Bootstrapping the GMM overidentification test under first-order underidentification, *CIRANO Working Paper* 2014s-25.
- Dufour J-M., Khalaf, L., and Kichian, M. (2013), Identification-robust analysis of DSGE and structural macroeconomic models, *Journal of Monetary Economics* 60, 340-350.
- Fanelli, L. (2012), Determinacy, indeterminacy and dynamic misspecification in linear rational expectations models, *Journal of Econometrics* 170, 153-163.
- Fanelli, L., Palomba, G. (2011), Simulation-based tests of forward-looking models under VAR learning Dynamics, *Journal of Applied Econometrics* 26, 762-782.
- Fernández-Villaverde, J., Rubio-Ramírez, J., Sargent, T., Watson, M. (2007), ABCs (and Ds) of understanding VARs, *American Economic Review* 97, 1021-1026.
- Fève, P., Matheron, J. and Sahuc, J-G. (2009), Minimum distance estimation and testing of DSGE model from structural VARs, *Oxford Bulletin of Economics and Statistics* 71, 883-894.
- Franchi, M., Vidotto, A. (2013), A check for finite order VAR representations of DSGE models, *Economic Letters* 120, 100-103.
- Franchi, M., Paruolo, P. (2015), Minimality of state space solutions of DSGE models and existence conditions for their VAR representation, *Computational Economics* 46, 613-626.
- Gonçalves, S. and L. Kilian (2004), Bootstrapping autoregressions with conditional heteroskedasticity of unknown form, *Journal of Econometrics* 123, 89-120.
- Guerron-Quintana P, Inoue A, Kilian L. (2013), Frequentist inference in weakly identified DSGE models, *Quantitative Economics* 4, 197-229.
- Hannan, E.J., Deistler, M. (1988), *The Statistical Theory of Linear Systems*, SIAM ed. Wiley, New York.
- Hansen, L.P. (2014), Nobel Lecture: Uncertainty outside and inside economic models, *Journal of Political Economy*, 122, 945-987.



- Hansen, L. P., Sargent, T. J. (1980), Formulating and estimating dynamic linear rational expectations models, *Journal of Economic Dynamics and Control* 2, 7-46.
- Hansen, L. P., Sargent, T. J. (1981), Linear rational expectations models for dynamically interrelated variables, in: R.E.Jr. Lucas and T. J. Sargent, (Eds.), *Rational Expectations and Econometric Practise*. University of Minnesota Press, Minneapolis, pp. 127-156.
- Horowitz, J.L. (2001), The bootstrap, *Handbook of Econometrics* Vol. 5, (eds.) Heckman, J.J and E. Leamer, Elsevier Science, pp. 3159-3226.
- Inoue, A., Kuo, C-H. and Rossi, B. (2015), Identifying the sources of model misspecification, Working Paper.
- Iskrev, N (2008), Evaluating the information matrix in linearized DSGE models, *Economic Letters* 99, 607-610.
- Iskrev, N (2010), Local identification in DSGE models, *Journal of Monetary Economics* 57, 189-202.
- Jarque, C. M, Bera, A. K. (1987), A test for normality of observations and regression residuals, *International Statistical Review* 55, 163-172.
- Kilian, L. (1998), Small-sample confidence intervals for impulse response functions, *Review of Economic and Statistics* 80, 218-230.
- Kleibergen, F., Mavroeidis, S. (2009), Weak instrument robust tests in GMM and the New Keynesian Phillips Curve, *Journal of Business and Economic Statistics* 27, 293-311.
- Komunjer, I., Ng, S. (2011), Dynamic identification of dynamic stochastic general equilibrium models, *Econometrica* 79, 1995-2032.
- Le, V.P.M, Meenagh, D., Minford, P. and Wickens, M. (2011), How much nominal rigidity is there in the US economy? Testing a new Keynesian DSGE model using indirect inference, *Journal of Economic Dynamics and Control* 35, 2078-2140.
- Leeper, E. M., Walker, T. B. and Yang, S.-C. S. (2013), Fiscal Foresight and Information Flows. *Econometrica* 81, 1115–1145.
- Liu, R.Y. (1988), Bootstrap procedures under some non i.i.d. models, *Annals of Statistics* 16, 1696–1708.
- Ljung, L., Caines, P.E. (1979), Asymptotic normality of predictor error estimators for approximate system models, *Stochastics* 3, 29-46.
- Lubik, T. A., Schorfheide, F. (2004), Testing for indeterminacy: an application to U.S. monetary policy, *American Economic Review* 94, 190-217.
- Lütkepohl, H. (2005), *New introduction to multiple time series*, Springer, Berlin.

- Mammen, E. (1993), Bootstrap and wild bootstrap for high dimensional linear models, *Annals of Statistics* 21, 255–285.
- Mankiw, N.G., Shapiro, M. D. (1986), Do we reject too often? *Economic Letters* 20, 139-145.
- Mavroeidis, S. (2005), Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve, *Journal of Money Credit and Banking* 37, 421-448.
- Mavroeidis, S. (2010), Monetary policy rules and macroeconomic stability: some new evidence, *American Economic Review* 100, 491-503.
- Moreira, M., Porter, J.R. and Suarez, G.A. (2004), Bootstrap and higher-order expansions validity when instruments may be weak, *NBER Working Paper* 302.
- Qu, Z. (2014), Inference in dynamic stochastic general equilibrium models with possible weak identification, *Quantitative Economics* 5, 457-494.
- Qu, Z., Tkachenko, D. (2012), Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models, *Quantitative Economics* 3, 95-132.
- Ravenna, F. (2007), Vector autoregressions and reduced form representations of DSGE models, *Journal of Monetary Economics* 54, 2048-2064.
- Schorfheide, F. (2010), Estimation and evaluation of DSGE models: Progress and challenges, NBER Working Paper n. 141.
- Shapiro, S. S., Wilk, M. B. (1965), An analysis of variance test for normality (complete samples), *Biometrika* 52, 591-611.
- Shorack, G.R., Wellner, J.A. (1986), *Empirical Processes with Applications to Statistics*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.
- Stoffer, D.S. and Wall, K. (1991), Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter, *Journal of the American Statistical Association* 86, 1024-1033.
- Stoffer, D.S. and Wall, K. (2004), Resampling in state space models, in *State Space and Unobserved Component Models: Theory and Applications*, (eds.) A.C. Harvey, S.J. Koopman and N. Shepard, Chap. 9, Cambridge University Press, Cambridge.
- Staiger, D., Stock, J.H. (1997), Instrumental variables regressions with weak instruments, *Econometrica* 65, 557-586.
- Stock, J.H., Wright, J. H. (2000), GMM with weak identification, *Econometrica* 68, 1055-1096.

- Watson, M. W. (1989), Recursive solution methods for dynamic linear rational expectations models, *Journal of Econometrics* 41, 65-89.
- White, H. (1982), Maximum likelihood estimation in misspecified models, *Econometrica* 50, 1-25.
- Wu, C.F.J. (1986), Jackknife, bootstrap, and other resampling methods, *Annals of Statistics* 14, 1261–1295.
- Zhan, Z. (2014), Detecting weak identification by bootstrap, mimeo.

TABLE 1. MC experiment. AS-DGP1

$\tau_0$	$\kappa_0$	$\psi_{1,0}$	$\psi_{2,0}$	$\beta_0$	$\rho_{r,0}$	$\rho_{g,0}$	$\rho_{z,0}$	$\sigma_{r,0}^2$	$\sigma_{g,0}^2$	$\sigma_{z,0}^2$
2	0.33	1.5	0.125	0.9975	0.75	0.95	0.90	0.04	0.36	0.09
T=100		Monte Carlo					Bootstrap (B=99)			
		$\hat{\theta}_{1,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{1,T}^*$	s.e.			
	$\kappa$	0.3311	0.0376	0.0399		0.3320	0.0381			
	$\rho_r$	0.7485	0.0178	0.0178		0.7471	0.0199			
	$\sigma_g^2$	0.3536	0.0503	0.0508		0.3512	0.0482			
		$Rej(LR_T) = 0.0786$				$Rej(LR_T^*) = 0.0208$				
Normality, JB		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.0625	0.0917	0.0833	0.0875	0.1667				
	$\hat{\rho}_{r,T}^*$	0.0417	0.0417	0.0500	0.0708	0.1458				
	$\hat{\sigma}_{g,T}^{2,*}$	0.0375	0.0292	0.0500	0.0500	0.1250				
Normality, JB (Proc.)		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.0875	0.0833	0.0917	0.1167	0.1667				
	$\hat{\rho}_{r,T}^*$	0.0583	0.0708	0.0750	0.0917	0.1458				
	$\hat{\sigma}_{g,T}^{2,*}$	0.0500	0.0750	0.0708	0.1000	0.1250				

NOTES: Results are based on M=500 Monte Carlo replications. Upper panel: ' $\hat{\theta}_{1,T}$ ' is the average of the ML estimator across Monte Carlo replications; 's.e.' indicates the standard errors computed across Monte Carlo replications; 's.e.(Hess.)' indicates the average of the Hessian-based standard errors across Monte Carlo replications; ' $\hat{\theta}_{1,T}^*$ ' is the average across Monte Carlo repetitions of the bootstrap ML estimators obtained as average of  $\hat{\theta}_{1,T:1}^*, \dots, \hat{\theta}_{1,T:B}^*$ ; the log-likelihood maximization under both  $H_0$  (with the CER) and  $H_1$  (without the CER) is obtained by the combining the Kalman filter with the 'CMA-ES' algorithm; 'Rej( $\cdot$ )' denotes rejection frequency across Monte Carlo replications,  $LR_T$  is the LR test for  $H_0$  against  $H_1$  and is computed using the 5% nominal significance level and  $LR_T^*$  is the (iid) bootstrap analog of  $LR_T$  based on  $B$  bootstrap repetitions. Lower panel: 'Normality, JB' reports rejection frequencies associated with Jarque and Bera's (1987) normality tests of the  $\tilde{B}$  bootstrap repetitions of the structural parameter estimators; 'Normality, JB' (Proc.) reports rejection frequencies associated with the procedure described in the Remark 4.7; the  $N := \text{int}(B/\tilde{B})$  sub-samples used for the procedure in Remark 4.7 are not overlapping; all normality tests are computed using the 5% nominal significance level.

TABLE 2. MC experiment. AS-DGP2

$\tau_0$	$\kappa_0$	$\psi_{1,0}$	$\psi_{2,0}$	$\beta_0$	$\rho_{r,0}$	$\rho_{g,0}$	$\rho_{z,0}$	$\sigma_{r,0}^2$	$\sigma_{g,0}^2$	$\sigma_{z,0}^2$
2	0.33	1.5	0.125	0.9975	0.75	0.95	0.90	0.04	0.36	0.09
T=100				Monte Carlo			Bootstrap (B=99)			
			$\hat{\theta}_{2,T}$	s.e.	s.e.(Hess.)	$\hat{\theta}_{2,T}^*$	s.e.			
	$\kappa$		0.3368	0.0587	0.0611	0.3408	0.0582			
	$\psi_1$		1.5121	0.1717	0.1205	1.5327	0.1644			
	$\rho_r$		0.7481	0.0260	0.0176	0.7480	0.0268			
	$\sigma_r^2$		0.0393	0.0061	0.0057	0.0382	0.0061			
	$\sigma_g^2$		0.3549	0.0517	0.0501	0.3456	0.0492			
	$\sigma_z^2$		0.0907	0.0242	0.0191	0.0925	0.0237			
			$Rej(LR_T) = 0.0785$			$Rej(LR_T^*) = 0.0331$				
Normality, JB		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.1064	0.1749	0.1868	0.2033	0.4397				
	$\hat{\psi}_{1,T}^*$	0.1277	0.2175	0.2766	0.3050	0.5816				
	$\hat{\rho}_{r,T}^*$	0.0662	0.0922	0.0946	0.1064	0.1395				
	$\hat{\sigma}_{r,T}^{2,*}$	0.0662	0.0804	0.0922	0.1348	0.1891				
	$\hat{\sigma}_{g,T}^{2,*}$	0.0638	0.0709	0.0898	0.0946	0.1608				
	$\hat{\sigma}_{z,T}^{2,*}$	0.0780	0.0969	0.1135	0.1537	0.2979				
Normality, JB (Proc.)		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.1253	0.1466	0.1749	0.2128	0.4397				
	$\hat{\psi}_{1,T}^*$	0.2009	0.3570	0.3948	0.4161	0.5816				
	$\hat{\rho}_{r,T}^*$	0.0686	0.0922	0.1135	0.0969	0.1395				
	$\hat{\sigma}_{r,T}^{2,*}$	0.0709	0.0851	0.0946	0.1158	0.1891				
	$\hat{\sigma}_{g,T}^{2,*}$	0.0686	0.0780	0.0946	0.1040	0.1608				
	$\hat{\sigma}_{z,T}^{2,*}$	0.0946	0.1418	0.1608	0.1820	0.2979				

NOTES: See Table 1.

TABLE 3. An and Shorfheide’s (2007) DSGE model on U.S. quarterly data

Sample 1984:Q2-2008Q3, T=98			Bootstrap		
model AS-M1	$\hat{\theta}_{1,T}$	s.e.(Hess.)	$\hat{\theta}_{1,T}^*$	s.e.	Normality
$\kappa$	0.1233	0.0177	0.1237	0.0210	$p$ -val.=0.2593(an)
$\rho_r$	0.8354	0.0174	0.7583	0.0301	$p$ -val.=0.2141(an)
$\sigma_g^2$	0.5862	0.1004	0.5681	0.1005	$p$ -val.=0.5000(an)
$LR_T = 176.11$ asym. $p$ -val.=0.0000			Bootstrap (iid) $p$ -val.=0.053		
			Bootstrap (wild) $p$ -val.=0.147		
model AS-M2	$\hat{\theta}_{2,T}$	s.e.(Hess.)	$\hat{\theta}_{2,T}^*$	s.e.	Normality
$\kappa$	0.6979	0.2160	0.6915	0.0368	$p$ -val.=0.0010(rn)
$\psi_1$	2.0000	0.2823	1.9403	0.1467	$p$ -val.=0.0018(rn)
$\rho_r$	0.7757	0.0169	0.6092	0.0767	$p$ -val.=0.1322(an)
$\sigma_r^2$	0.0291	0.0044	0.0646	0.0293	$p$ -val.=0.0328(an)
$\sigma_g^2$	0.3113	0.0479	0.3150	0.0536	$p$ -val.=0.5000(an)
$\sigma_z^2$	0.0163	0.0033	0.0221	0.0087	$p$ -val.=0.0215(rn)
$LR_T = 83.25$ asym. $p$ -val.=0.0000			Bootstrap (iid) $p$ -val.=0.007		
			Bootstrap (wild) $p$ -val.=0.112		

NOTES: The log-likelihood maximization under both  $H_0$  (with the CER) and  $H_1$  (without the CER) is obtained by the combining the Kalman filter with the ‘CMA-ES’ algorithm with the following bounds on the optimization parameter space: [0.04, 0.7] for  $\kappa$ , [0.5, 0.95] for  $\rho_r$  and [0.5, 2] for  $\psi_1$ . The column ‘s.e.(Hess.)’ summarizes the standard errors computed using the inverse of the estimated Hessian matrix; the columns under ‘Bootstrap’ report averages across  $B=999$  (iid) bootstrap repetitions; asymptotic p-values associated with the QLR tests for the CER are taken from the  $\chi_k^2$  distribution with  $k=10=13-3$  for model AS-M1 and  $k=7=13-6$  for model AS-M2, where  $13=\dim(\phi)$ , and  $\phi$  is the vector of parameters of the state space representation of the DSGE model under  $H_1$ ; the bootstrap (iid) p-values associated with the QLR tests are computed using  $B=999$  repetitions generated under  $H_0$  and the algorithm of Section 4.1; the p-value calculated with the wild-bootstrap is obtained as described in the Remark 4.3 using  $w_t^* \sim N(0,1)$ ; the p-values under ‘Normality’ refer to Jarque and Bera’s (1987) normality test calculated on  $\tilde{B}=20$  bootstrap repetitions; ‘rn’ means that the procedure summarized in Remark 4.7 rejects the null of normality, ‘an’ means that the procedure accepts the null of normality.

# TECHNICAL SUPPLEMENT TO 'BOOTSTRAPPING DSGE MODELS'

BY GIOVANNI ANGELINI<sup>†</sup>, GIUSEPPE CAVALIERE<sup>†</sup> AND LUCA FANELLI<sup>†\*</sup>

August 29, 2016

## TS.1 INTRODUCTION

This Technical Supplement develops/expands a number of topics discussed in Angelini, Cavaliere and Fanelli (2016), ACF hereafter. It is organized as follows.

In Section TS.2 we present the relationship between the structural form of the DSGE model, its rational expectations determinate solution and the associated 'ABCD' form. In Section TS.3 we illustrate the relationship between the coefficients that characterize the state space representation associated with the DSGE model without the CER imposed and the structural parameters. Section TS.4 is devoted to some estimation issues. In particular, in Section TS.4.1 we discuss QML estimation of the DSGE model with and without the CER imposed, while in Section TS.4.2 we review the regularity conditions that permit standard inference in strongly identified DSGE models. In Section TS.5 we provide further Monte Carlo results based on the ARMA(1,1) process (Section TS.5.1), on Andrews and Mikusheva's (2015) DSGE model (Section TS.5.2), and we also include some additional tables related to An and Schorfheide's (2007) DSGE model (Section TS.5.3) analyzed in ACF. Finally, further empirical results related to the empirical illustration in ACF are given in Section TS.6.

## TS.2 STRUCTURAL MODEL, RATIONAL EXPECTATIONS SOLUTION AND STATE SPACE REPRESENTATION

A common starting point in the econometric analysis of DSGE models is the so-called 'ABCD' state space representation, see Fernández-Villaverde *et al.* (2007), reported in eq.s (10)-(11) in ACF. It is not always clear, however, which is the connections between the structural form, its rational expectations solution and the ABCD form. We start from the structural representation of the DSGE model, i.e. from the system of log-linearized structural Euler equations that describe the economy, and then discuss how to obtain the associated ABCD state space form.

Let  $Z_t := (Z_{1,t}, Z_{2,t}, \dots, Z_{n_z,t})'$  be a  $n_z \times 1$  vector of endogenous possibly unobserved variables at time  $t$ , which can be interpreted as deviation from steady state values. We assume that, after log-linearization and, for  $t = 1, \dots, T$ , the structural form reads as

$$\Gamma_0 Z_t = \Gamma_f E_t Z_{t+1} + \Gamma_b Z_{t-1} + \Pi \eta_t \quad (\text{TS.1})$$

$$\eta_t = R \eta_{t-1} + \omega_t \quad , \quad \omega_t \sim \text{WN}(0, \Sigma_\omega). \quad (\text{TS.2})$$

In eq.s (TS.1)-(TS.2),  $\Gamma_i := \Gamma_i(\theta)$ ,  $i = 0, f, b$  are  $n_z \times n_z$  matrices whose elements depend on the vector of structural parameters  $\theta$ ,  $\Pi := \Pi(\theta)$  is an  $n_z \times n_\omega$  matrix of full-column rank ( $n_\omega \leq n_z$ ), whose elements may depend on  $\theta$  but which essentially selects the shocks that enter the equations,  $\eta_t$  is a  $n_\omega \times 1$  vector of autoregressive disturbances,  $R := R(\theta)$  is an  $n_\omega \times n_\omega$  stable diagonal matrix,  $\omega_t$  is the  $n_\omega \times 1$  vector of white noise structural shocks with covariance matrix  $\Sigma_\omega := \Sigma_\omega(\theta)$ .  $\omega_t$  is adapted to the sigma-field  $\mathcal{F}_t$ , where  $\mathcal{F}_t$  is the agents' information set at time  $t$  and  $E_t Z_{t+1} := E(Z_{t+1} | \mathcal{F}_t)$ . The term  $\omega_t$  will be referred to as the vector of fundamental structural shocks, and its covariance matrix  $\Sigma_\omega$  can be either diagonal or non-diagonal. The initial condition  $Z_0$  are treated as given. Finally, the true value of  $\theta$  is denoted with  $\theta_0$  and is assumed to be an interior point of the determinacy region  $\mathcal{P}^D$  of the compact parameter space  $\mathcal{P}$ .

The multivariate linear(ized) rational expectations model described by system (TS.1)-(TS.2) nests a large class of (log-linearized) DSGE models currently used in policy and business cycle analysis. A convenient summary of system (TS.1)-(TS.2) can be obtained by re-defining the vector of state variables as  $Z_t^\circ := (Z_t', \eta_t')'$ , and considering the  $(n_z + n_\omega)$ -dimensional system

$$\begin{aligned} \begin{pmatrix} \Gamma_0 & -\Pi \\ 0_{n_\omega \times n_z} & I_{n_\omega} \\ \Gamma_0^\circ & \end{pmatrix} \begin{pmatrix} Z_t \\ \eta_t \\ Z_t^\circ \end{pmatrix} &= \begin{pmatrix} \Gamma_f & 0_{n_z \times n_\omega} \\ 0_{n_\omega \times n_z} & 0_{n_\omega \times n_\omega} \\ \Gamma_f^\circ & \end{pmatrix} \begin{pmatrix} E_t Z_{t+1} \\ E_t \eta_{t+1} \\ E_t Z_{t+1}^\circ \end{pmatrix} \\ &+ \begin{pmatrix} \Gamma_b & 0_{n_z \times n_\omega} \\ 0_{n_\omega \times n_z} & R \\ \Gamma_b^\circ & \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ \eta_{t-1} \\ Z_{t-1}^\circ \end{pmatrix} + \begin{pmatrix} 0_{n_z \times n_\omega} \\ I_{n_\omega} \\ F^\circ \end{pmatrix} \omega_t \end{aligned}$$

that we compact in the expression

$$\Gamma_0^\circ Z_t^\circ = \Gamma_f^\circ E_t Z_{t+1}^\circ + \Gamma_b^\circ Z_{t-1}^\circ + F^\circ \omega_t. \quad (\text{TS.3})$$

The definition of the matrices  $\Gamma_0^\circ$ ,  $\Gamma_f^\circ$ ,  $\Gamma_b^\circ$  and  $F^\circ$  in system (TS.3) is obvious.

A rational expectations solution to system (TS.3) is any stochastic process  $\{Z_t^{\circ s}(\theta)\}_{t=0}^\infty$  such that, for  $\theta \in \mathcal{P}_\theta$  and for fixed initial conditions, the quantity  $E_t Z_{t+1}^{\circ s}(\theta) = E(Z_{t+1}^{\circ s}(\theta) | \mathcal{F}_t)$  exists and if  $Z_{t+1}^{\circ s}(\theta)$  is substituted for  $Z_t^\circ$  ( $Z_t^{\circ s}(\theta) = Z_t^\circ$ ) into the structural equations, the model is verified for each  $t$ . We use the notation  $\{Z_t^{\circ s}(\theta)\}_{t=0}^\infty$  to remark the dependence of the solution on the values of the structural parameters.

Under Assumption 1 of ACF, the unique stable solution associated with system (TS.3) can be represented in the form

$$Z_t^\circ = \Psi^\circ(\phi_\theta) Z_{t-1}^\circ + N^\circ(\phi_\theta) \omega_t, \quad t = 1, \dots, T \quad (\text{TS.4})$$

where the  $(n_z + n_\omega) \times (n_z + n_\omega)$  matrix  $\Psi^\circ := \Psi^\circ(\phi_\theta)$  is stable and solves the second-order quadratic matrix equation

$$\Gamma_f^\circ (\Psi^\circ)^2 - \Gamma_0^\circ \Psi^\circ + \Gamma_b^\circ = 0_{(n_z + n_\omega) \times (n_z + n_\omega)}, \quad (\text{TS.5})$$

while the  $(n_z + n_\omega) \times n_\omega$  matrix  $N^\circ := N^\circ(\phi_\theta)$  satisfies

$$(\Gamma_0^\circ - \Gamma_f^\circ \Psi^\circ) N^\circ = F^\circ. \quad (\text{TS.6})$$



The relationships in eqs. (TS.5)-(TS.6) can be easily obtained using e.g. Binder and Pesaran's (1995) solution method, see also Uhlig (1999). In eq. (TS.4), we use the notation  $\phi_\theta = g(\theta)$ , where  $g(\cdot)$  is a nonlinear vector function, to remark the fact that there exists a set of 'reduced form' coefficients, collected in the vector  $\phi_\theta$ , that depend nonlinearly on the structural parameters  $\theta$  through the CER in eqs. (TS.5)-(TS.6). Obviously,  $\phi_{\theta_0} = g(\theta_0)$ . We come back on this in Section TS.3. The  $\Psi^\circ$  matrix that solves eq. (TS.5) has the following structure

$$\Psi^\circ := \begin{pmatrix} \Psi_{11}^\circ & \Psi_{12}^\circ \\ 0_{n_\omega \times n_z} & R \end{pmatrix}$$

and  $\Psi_{11}^\circ$  and  $\Psi_{12}^\circ$  are  $n_z \times n_z$  and  $n_z \times n_\omega$ , respectively.

Let  $y_t := (y_{1,t}, y_{2,t}, \dots, y_{n_y,t})'$  be the  $n_y \times 1$  vector of (demeaned) observed variables. The solution in eq. (TS.4) can be viewed as a state system with associated measurement system

$$\underset{n_y \times 1}{y_t} = \underset{n_y \times (n_z + n_\omega)}{H} \underset{(n_z + n_\omega) \times 1}{Z_t^\circ} \quad (\text{TS.7})$$

where  $H$  plays the role of a  $n_y \times (n_z + n_\omega)$  selection matrix that picks out the observables from  $Z_t^\circ$ . Although eq. (TS.7) is general enough, we might consider a more general formulation, given by

$$\underset{n_y \times 1}{y_t} = \underset{n_y \times (n_z + n_\omega)}{H} \underset{(n_z + n_\omega) \times 1}{Z_t^\circ} + \underset{n_y \times n_v}{J} \underset{n_v \times 1}{v_t}. \quad (\text{TS.8})$$

In this specification,  $v_t$  is a  $n_v \times 1$  ( $n_v \leq n_y$ ) vector of measurement errors, assumed white noise with covariance matrix  $\Sigma_v$ , and  $J$  is a  $n_y \times n_v$  selection matrix. The vector  $v_t$  can be interpreted as an 'additional' source of random shocks (errors) that impinge on  $y_t$ , and is generally interpreted as a measurement error, but it can also involve elements of  $\omega_t$  itself (see e.g. the measurement equations in An and Schorfheide, 2007). Typically, in singular systems, i.e. models for which  $n_y > n_\omega$ , a number  $n_v = n_y - n_\omega$  of measurement errors is artificially added to the system such that the  $n_u$ -dimensional ( $n_u := n_\omega + n_v$ ) vector  $u_t := (\omega_t', v_t')$  is the new vector of models' innovations, and the equality  $n_y = n_u$  is obtained. In the derivations that follow, we stick to the measurement system in eq. (TS.7) (our approach can easily be generalized to the case where  $\omega_t$  is replaced with  $u_t$ ).

The solution in eq. (TS.4) obtained under Assumption 1 is not generally expressed in 'minimal' form, meaning that  $Z_t^\circ$  might contain more states than strictly necessary to fully characterize the overall dynamics of the system. In principle, a more parsimonious state space representation can be achieved by selecting the relevant states because any non-minimal form can be manipulated such that eventually a minimal representation is achieved. Extensive discussions and examples of how this can be done in practice may be found in e.g. Komunjer and Ng (2011) and Guerron-Quintana *et al.* (2013).

A natural criterion to obtain a minimal representation from system (TS.4) is provided by the columns of the matrix  $\Psi^\circ$ : zero columns correspond to non-minimal states, see e.g. Komunjer and Ng (2011). Thus, dropping states in  $Z_t^\circ$  associated with zero columns of  $\Psi^\circ$  is a necessary, albeit not sufficient, operational step to obtain a minimal

representation. Let  $\mathcal{P}_m$  be the  $(n_z + n_\omega)$ -dimensional permutation matrix such that  $\mathcal{P}_m Z_t^\circ = (Z'_{m,t}, Z'_{r,t})$ , where  $Z'_{m,t}$  is the sub-vector of states which are associated with non zero-columns of  $\Psi^\circ$ , and  $Z'_{r,t}$  is the sub-vector of states associated with zero-columns of  $\Psi^\circ$ . Using  $\mathcal{P}_m$ , we can re-parameterize system (TS.4) in the form

$$\mathcal{P}_m Z_t^\circ = \mathcal{P}_m \Psi^\circ(\phi_\theta) \mathcal{P}_m^{-1} \mathcal{P}_m Z_{t-1}^\circ + \mathcal{P}_m N^\circ(\phi_\theta) \omega_t$$

which corresponds to the representation

$$\begin{pmatrix} Z_{m,t} \\ Z_{r,t} \end{pmatrix} = \begin{pmatrix} \Psi_m(\phi_\theta) & 0_{n_m \times (n_z + n_\omega - n_m)} \\ \Psi_r(\phi_\theta) & 0_{(n_z + n_\omega - n_m) \times (n_z + n_\omega - n_m)} \end{pmatrix} \begin{pmatrix} Z_{m,t-1} \\ Z_{r,t-1} \end{pmatrix} + \begin{pmatrix} N_m(\phi_\theta) \\ N_r(\phi_\theta) \end{pmatrix} \omega_t \quad (\text{TS.9})$$

where  $Z_{m,t}$  is the  $n_m \times 1$  vector of potentially minimal state variables and  $Z_{r,t}$  is the  $(n_z + n_\omega - n_m) \times 1$  vector of ‘redundant’ state variables. Observe that  $n_m \leq n_z + n_\omega$ , and that  $\Psi_m := \Psi_m(\phi_\theta)$ ,  $\Psi_r := \Psi_r(\phi_\theta)$ ,  $N_m := N_m(\phi_\theta)$  and  $N_r := N_r(\phi_\theta)$  are matrices of suitable dimensions derived from  $\Psi^\circ$  and  $N^\circ$  by properly rearranging states.

From the first set of  $n_m$  equations of system (TS.9) we obtain

$$Z_{m,t} = \Psi_m(\phi_\theta) Z_{m,t-1} + N_m(\phi_\theta) \omega_t. \quad (\text{TS.10})$$

Using  $\mathcal{P}_m$  and eq. (TS.7),<sup>1</sup> the measurement system reads

$$\begin{aligned} y_t &= H \mathcal{P}_m^{-1} \mathcal{P}_m Z_t^\circ = H \mathcal{P}_m^{-1} \begin{pmatrix} Z_{m,t} \\ Z_{r,t} \end{pmatrix} \\ &= (H_m : H_r) \begin{pmatrix} Z_{m,t} \\ Z_{r,t} \end{pmatrix} = H_m Z_{m,t} + H_r Z_{r,t} \\ &= \underbrace{(H_m \Psi_m + H_r \Psi_r)}_{n_y \times n_m} \underbrace{Z_{m,t-1}}_{n_m \times 1} + \underbrace{(H_m N_m + H_r N_r)}_{n_y \times n_\omega} \underbrace{\omega_t}_{n_\omega \times 1} \end{aligned} \quad (\text{TS.11})$$

where the selection matrices  $H_m$  and  $H_r$  satisfy the condition  $H \mathcal{P}_m^{-1} = (H_m : H_r)$ . By coupling system (TS.10) and system (TS.11) and re-naming matrices as follows:  $A(\phi_\theta) := \Psi_m(\phi_\theta)$ ,  $B(\phi_\theta) := N_m(\phi_\theta)$ ,  $C(\phi_\theta) := (H_m \Psi_m + H_r \Psi_r)$ ,  $D(\theta) := (H_m N_m + H_r N_r)$ , we obtain the so-called ABCD form:

$$Z_{m,t} = A(\phi_\theta) Z_{m,t-1} + B(\phi_\theta) \omega_t \quad (\text{TS.12})$$

$n_m \times 1$        $n_m \times n_m$     $n_m \times 1$        $n_m \times n_\omega$     $n_\omega \times 1$

$$y_t = C(\phi_\theta) Z_{m,t-1} + D(\phi_\theta) \omega_t \quad (\text{TS.13})$$

$n_y \times 1$        $n_y \times n_m$     $n_m \times 1$        $n_y \times n_\omega$     $n_\omega \times 1$

which is the candidate minimal state space representation associated with the DSGE model discussed in eq.s (10)-(11) of Section 2 of ACF.

### TS.3 MAPPING BETWEEN STATE SPACE PARAMETERS AND STRUCTURAL PARAMETERS UNDER THE CER

Given the ABCD form in eq.s (TS.12)-(TS.13) and Assumptions 1-3 of ACF, the ‘non-redundant’ CER in eq. (TS.5) can be re-formulated as follows. Let  $\Psi_m^*$  the block

<sup>1</sup>A equivalent but more involved derivation follows if eq. (TS.8) is used in place of eq. (TS.7).

matrix that characterizes system (TS.9). The relationship between  $\Psi_m^*$  and  $\Psi^*$  is given by  $\Psi_m^* = \mathcal{P}_m \Psi^* \mathcal{P}_m^{-1}$ , hence the restrictions entailed by the quadratic matrix equation in eq. (TS.5) can be re-written as

$$\begin{aligned} & \Gamma_f^* (\mathcal{P}_m^{-1} \Psi_m^* \mathcal{P}_m)^2 - \Gamma_0^* (\mathcal{P}_m^{-1} \Psi_m^* \mathcal{P}_m) + \Gamma_b^* \\ = & \Gamma_f^* \mathcal{P}_m^{-1} (\Psi_m^*)^2 \mathcal{P}_m - \Gamma_0^* \mathcal{P}_m^{-1} \Psi_m^* \mathcal{P}_m + \Gamma_b^* = 0_{(n_z+n_\omega) \times (n_z+n_\omega)} \end{aligned}$$

which are equivalent to

$$\Gamma_f^{**} (\Psi_m^*)^2 - \Gamma_0^* \Psi_m^* + \Gamma_b^{**} = 0_{(n_z+n_\omega) \times (n_z+n_\omega)} \quad (\text{TS.14})$$

where  $\Gamma_f^{**} := \Gamma_f^* \mathcal{P}_m^{-1}$ ,  $\Gamma_0^{**} := \Gamma_0^* \mathcal{P}_m^{-1}$  and  $\Gamma_b^{**} := \Gamma_b^* \mathcal{P}_m^{-1}$ . The expression in eq. (TS.14) gives rise to both trivial identities (i.e. zeros on the left- and right-hand side of the system) and actual constraints. It is convenient to summarize the constraints (not including trivial identities) featured by eq. (TS.14) as

$$f(\theta, \phi) = 0_{\dim(\phi) \times 1} \quad (\text{TS.15})$$

where  $f(\cdot, \cdot)$  is a nonlinear twice-differentiable vector function, and  $\phi$  is the vector that contains all non-zero elements of the matrix  $\Psi_m^*$  that contribute to the likelihood of the state space representation of the model without the CER;  $\dim(\phi) < n_z + n_\omega$ . If the CER are valid at the true parameters points and

$$\det \left( \left. \frac{\partial f(\theta, \phi)}{\partial \phi'} \right|_{(\theta, \phi) = (\theta_0, \phi_0)} \right) \neq 0,$$

by the implicit function theorem  $\phi = g(\theta)$ , where  $g(\cdot)$  is the nonlinear differentiable function already mentioned in Section TS.2. Thus,  $\phi$  can be qualified as the vector that contains all non-zero elements of the matrix  $\Psi_m^*$  before the CER are imposed. Once the CER are imposed, we obtain  $\phi_\theta = g(\theta)$ , where  $g(\cdot)$  is such that

$$\frac{\partial \phi_\theta}{\partial \theta'} \Big|_{\dim(\phi) \times \dim(\theta)} = - \left( \frac{\partial f(\theta, \phi)}{\partial \phi} \Big|_{\dim(\phi) \times \dim(\phi)} \right)^{-1} \times \left( \frac{\partial f(\theta, \phi)}{\partial \theta'} \Big|_{\dim(\phi) \times \dim(\theta)} \right). \quad (\text{TS.16})$$

Hence,  $\phi_\theta$  ( $\dim(\phi_\theta) = \dim(\phi)$ ), can be qualified as the vector that contains all non-zero elements of the matrix  $\Psi_m^*$  after the CER are imposed.

Iskrev (2008) notices that the Jacobian in eq. (TS.16) can be used to derive analytic standard errors for the parameters of the estimated DSGE model, provided  $\hat{\phi}_T$  and a consistent estimate of the information matrix associated with the state space form is obtained. Indeed, let

$$\tilde{I}_{\hat{\phi}_T, T}^{OP} := \frac{1}{T} \sum_{t=1}^T \frac{\partial l(y_t | \mathcal{F}_{t-1}^y; \phi)}{\partial \phi} \times \frac{\partial l(y_t | \mathcal{F}_{t-1}^y; \phi)}{\partial \phi'} \Big|_{\phi = \hat{\phi}_T}$$

be the ‘incremental observed information’ estimate of the information matrix associated with the state space form that does not incorporate the CER, and  $l(y_t | \mathcal{F}_{t-1}^y; \phi)$  the log-likelihood associated with observations at time  $t$ . Using the chain rule we have:

$$\frac{\partial l(y_t | \mathcal{F}_{t-1}^y; g(\theta))}{\partial \theta} = \left( \frac{\partial g(\theta)}{\partial \theta'} \right)' \Big|_{\dim(\theta) \times \dim(\phi)} \times \frac{\partial l(y_t | \mathcal{F}_{t-1}^y; \phi)}{\partial \phi} \Big|_{\dim(\phi) \times 1}$$

so that

$$\begin{aligned}\tilde{\mathcal{I}}_{\hat{\theta}_T, T}^{OP} &= \frac{1}{T} \sum_{t=1}^T \left( \left( \frac{\partial g(\theta)}{\partial \theta'} \right)' \times \frac{\partial l(y_t | \mathcal{F}_{t-1}^y; \phi)}{\partial \phi} \right) \times \left( \frac{\partial l(y_t | \mathcal{F}_{t-1}^y; \phi)}{\partial \phi} \times \left( \frac{\partial g(\theta)}{\partial \theta'} \right) \right) \Bigg|_{\theta=\hat{\theta}_T, \phi=\hat{\phi}_T} \\ &= \left( \frac{\partial g(\theta)}{\partial \theta'} \right)_{\theta=\hat{\theta}_T} \tilde{\mathcal{I}}_{\hat{\phi}_T, T}^{OP} \left( \frac{\partial g(\theta)}{\partial \theta'} \right)_{\theta=\hat{\theta}_T}\end{aligned}$$

where this matrix is positive definite by construction.

## TS.4 ESTIMATION ISSUES

In this section we discuss some technicalities related to the maximum likelihood estimation of the DSGE model considered in ACF (section TS.4.1). We also review the regularity conditions that permit standard inference (Section TS.4.2),

### TS.4.1 DERIVATION OF THE QML ESTIMATOR OF THE DSGE MODEL

Estimation of the DSGE model under the null of the CER ( $H_0$ ) and without the CER ( $H_1$ ) is a necessary preliminary step to the computation of the QLR test for the CER discussed in Section 3 of ACF.

For convenience, we start from the ABCD representation:

$$\begin{matrix} Z_{m,t} \\ n_m \times 1 \end{matrix} = \begin{matrix} A(\phi_\theta) & B(\phi_\theta) \\ n_m \times n_m & n_m \times n_\omega \end{matrix} \begin{matrix} Z_{m,t-1} \\ n_m \times 1 \end{matrix} + \begin{matrix} \omega_t \\ n_\omega \times 1 \end{matrix} \quad (\text{TS.17})$$

$$\begin{matrix} y_t \\ n_y \times 1 \end{matrix} = \begin{matrix} C(\phi_\theta) & D(\phi_\theta) \\ n_y \times n_m & n_y \times n_\omega \end{matrix} \begin{matrix} Z_{m,t-1} \\ n_m \times 1 \end{matrix} + \begin{matrix} \omega_t \\ n_\omega \times 1 \end{matrix} \quad (\text{TS.18})$$

and consider Assumptions 1-3 and the other auxiliary hypotheses discussed in ACF. Using the invertibility of  $A(\phi_\theta)$ , the system can be written in the form

$$Z_{m,t} = A(\phi_\theta)Z_{m,t-1} + B(\phi_\theta)\omega_t \quad (\text{TS.19})$$

$$y_t = H^+(\phi_\theta)Z_{m,t} + J(\phi_\theta)\omega_t \quad (\text{TS.20})$$

where  $H^+(\phi_\theta) := C(\phi_\theta)A(\phi_\theta)^{-1}$  and  $J(\phi_\theta) := D(\phi_\theta) - C(\phi_\theta)A(\phi_\theta)^{-1}B(\phi_\theta)$ . For ease of notation, we suppress temporarily the dependence of the matrices appearing in system (TS.19)-(TS.20) on  $\phi_\theta = g(\theta)$ , and simply replace these matrices with  $A$ ,  $B$ ,  $H^+$  and  $J$ . Define  $\hat{Z}_{m,t|t-1} := E(Z_{m,t} | \mathcal{F}_{t-1}^y)$  and let  $\hat{y}_{t|t-1} := E(y_t | \mathcal{F}_{t-1}^y) = H^+ \hat{Z}_{m,t|t-1}$  be the best linear predictor of  $y_t$  based on the data in  $\mathcal{F}_{t-1}^y$ , where  $\mathcal{F}_t^y := \sigma(y_t, \dots, y_1) \subset \mathcal{F}_t$  is the information set based on the observable variables up to time  $t$ . Observe that  $E(\omega_t | \mathcal{F}_{t-1}^y) = 0_{n_\omega \times 1}$ .

The innovation form representation (Anderson and Moore, 1979) associated with system (TS.19)-(TS.20) can be written in the form

$$\hat{Z}_{m,t+1|t} = A \hat{Z}_{m,t|t-1} + K_t \epsilon_t^0(\theta) \quad (\text{TS.21})$$

$$y_t = H^+ \hat{Z}_{m,t|t-1} + \epsilon_t^0(\theta) \quad (\text{TS.22})$$

where  $K_t = K_t(\phi_\theta)$  is the Kalman gain and

$$\epsilon_t^0(\theta) := y_t - H^+ \hat{Z}_{m,t|t-1} \quad (\text{TS.23})$$

are the innovation residuals with covariance matrix

$$\Sigma_{\epsilon^0,t}(\theta) = H^+ P_{t|t-1}^0 H^{+'} + J \Sigma_\omega J' \quad (\text{TS.24})$$

and  $P_{t|t-1}^0 := E((Z_{m,t} - \hat{Z}_{m,t|t-1})(Z_{m,t} - \hat{Z}_{m,t|t-1})' | \mathcal{F}_{t-1}^y)$ ,  $P_{1|0}^0$  being given. We use the superscript ‘0’ for  $\epsilon_t^0$  in eq. (TS.23) and  $\Sigma_{\epsilon^0,t}$  and  $P_{t|t-1}^0$  in eq. (TS.24) to remark that the representation in eq.s (TS.21)-(TS.22) is obtained under the null  $H_0$  which imposes the CER. Imposing the normality of  $\epsilon_t^0$  in eq. (TS.23), i.e.

$$y_t | \mathcal{F}_{t-1}^y \sim N(H^+ \hat{Z}_{m,t|t-1}, \Sigma_{\epsilon^0,t}(\theta))$$

the estimation of  $\theta$  can be accomplished via Gaussian maximum likelihood estimation. Let  $\ell_T(g(\theta))$  be the Gaussian log-likelihood function associated with the state space model in eq.s (TS.21)-(TS.22). The essential part of the log-likelihood  $\ell_T(g(\theta))$ , denoted for simplicity by  $\ell_{o,T}(g(\theta)) := \sum_{t=1}^T l(y_t | \mathcal{F}_{t-1}^y; \theta)$ , is given by

$$\ell_{o,T}(\theta) : = - \sum_{t=1}^T \ell_t(\theta) \quad (\text{TS.25})$$

$$\ell_t(\theta) : = l(y_t | \mathcal{F}_{t-1}^y; \theta) = \{\log \det(\Sigma_{\epsilon^0,t}(\theta)) + \epsilon_t^0(\theta)' \Sigma_{\epsilon^0,t}(\theta)^{-1} \epsilon_t^0(\theta)\}$$

where  $\epsilon_t^0(\theta)$  and  $\Sigma_{\epsilon^0,t}(\theta)$  are defined above and depend on  $\theta$  through their dependence on  $A(\phi_\theta)$ ,  $B(\phi_\theta)$ ,  $H^+(\phi_\theta)$  and  $J(\phi_\theta)$ . Given  $\ell_{o,T}(\theta)$  in eq. (TS.25), the QML estimator of  $\theta$  solves

$$\hat{\theta}_T := \arg \max_{\theta \in \mathcal{P}^D} \ell_{o,T}(g(\theta)) \quad (\text{TS.26})$$

and can be computed by combining the Kalman filter with numerical optimization methods, see e.g. Andreasen (2010). The QML estimator of  $\phi_\theta$  is obtained from  $\hat{\phi}_{\theta,T} = g(\hat{\theta}_T)$ .

To estimate  $\phi$ , we can consider analogs of systems (TS.19)-(TS.20) and (TS.21)-(TS.22), with the matrices  $A$ ,  $K$ ,  $H^+$  now depending on  $\phi$ , and

$$\epsilon_t^u(\phi) y_t - H^+(\phi) \hat{Z}_{m,t|t-1}$$

$$\Sigma_{\epsilon^u,t}(\phi) = H^+(\phi) P_{t|t-1}^u H^+(\phi)' + J(\phi) \Sigma_\omega J(\phi)',$$

for suitably given  $P_{t|t-1}^u$ . We use the superscript ‘u’ for  $\epsilon_t^u(\phi)$ ,  $\Sigma_{\epsilon^u,t}(\phi)$  and  $P_{t|t-1}^u$  to remark that the innovation form depends on the vector  $\phi$  under  $H_1$ . The QML estimator of  $\phi$  is therefore obtained from

$$\hat{\phi}_T := \arg \max_{\phi \in \mathcal{P}_\phi} \ell_{o,T}(\phi), \quad \ell_{o,T}(\phi) = - \sum_{t=1}^T \left\{ \log \det(\Sigma_{\epsilon^u,t}(\phi)) + \epsilon_t^u(\phi)' \Sigma_{\epsilon^u,t}(\phi)^{-1} \epsilon_t^u(\phi) \right\}. \quad (\text{TS.27})$$

## TS.4.2 REGULARITY CONDITIONS FOR INFERENCE

In this section we review the regularity conditions that permit standard inference, hence strong identification, when the DSGE model is estimated by QML. We also stress the conditions that fail in weakly identified DSGE models and invalidate the Gaussian asymptotic approximation. The general setup is the same as in Section 2 of ACF, with Assumptions 1-3 and  $\theta_0$  replaced with the ‘pseudo-true’ value of the structural parameters  $\theta^\dagger$  (Section 3.2 of ACF). When the determinate solution of the DSGE model belongs to the data generating process, namely the CER under  $H_0$  are true,  $\theta^\dagger = \theta_0$ .

### CONSISTENCY OF THE QML ESTIMATOR

Assume the following conditions hold:

- (TS-A.i)  $\mathcal{P}^D$  is compact;
- (TS-A.ii)  $\theta^\dagger$  is the unique maximizer of  $Q_0(\theta)$  in a neighborhood  $\mathcal{N}_{\theta^\dagger}$  of  $\theta^\dagger$ ;
- (TS-A.iii)  $Q_0(\theta)$  is continuous in  $\theta$ ;
- (TS-A.iv)  $\hat{Q}_T(\theta)$  converge uniformly in probability to  $Q_0(\theta)$ , i.e.  $\sup_{\theta \in \mathcal{P}^D} \left| \hat{Q}_T(\theta) - Q_0(\theta) \right| \rightarrow_p 0$ .<sup>2</sup>

Then

$$\arg \max_{\theta \in \mathcal{P}^D} \hat{Q}_T(\theta) =: \hat{\theta}_T \rightarrow_p \theta^\dagger := \arg \max_{\theta \in \mathcal{P}^D} Q_0(\theta). \quad (\text{TS.28})$$

The proof of this standard result may be found, among others, in Wooldridge (1994).

In our framework, condition (TS-A.i) is covered by Assumption 1 in ACF. Condition (TS-A.ii) is consistent with the local identification condition postulated for the DSGE model in Assumptions 3 in ACF. Conditions (TS-A.iii) and (TS-A.iv) are standard regularity conditions for the consistency of the QML estimator  $\hat{\theta}_T$ . (TS-A.iii) is satisfied in our setup. For covariance stationary and ergodic processes, sufficient condition for (TS-A.iv) is that  $|\ell_t(\theta)| \leq b(y_t)$  for all  $\theta \in \mathcal{P}^D$ , for some function  $b(\cdot) \in \mathbb{R}^+$  such that  $E(b(y_t)) < \infty$ , see e.g. Theorem 4.1 in Wooldridge (1994). It is the lack of uniform convergence in (TS-A.iv) that may fail in weakly identified DSGE models.

### ASYMPTOTIC NORMALITY OF THE QML ESTIMATOR

Assume that in addition to (TS-A.i)-(TS-A.iv), the following conditions hold:

- (TS-A.v)  $\theta^\dagger$  is in  $\text{int}(\mathcal{P}^D)$ , where  $\text{int}(\cdot)$  indicates the interior of  $\mathcal{P}^D$ ;
- (TS-A.vi)  $T^{1/2} \nabla_\theta \hat{Q}_T(\theta^\dagger) = T^{-1/2} \sum_{t=1}^T \nabla_\theta \ell_t(\theta^\dagger)$  satisfies the central limit theorem (CLT):

$$T^{-1/2} \sum_{t=1}^T \nabla_\theta \ell_t(\theta^\dagger) \xrightarrow{d} N(0_{\dim(\theta) \times 1}, \mathcal{B}) \quad , \quad \mathcal{B} := \mathcal{B}(\theta^\dagger) := \lim_{T \rightarrow \infty} \text{Var} \left( T^{-1/2} \sum_{t=1}^T \nabla_\theta \ell_t(\theta^\dagger) \right);$$

(TS-A.vii)

$$\sup_{\theta \in \mathcal{N}_{\theta^\dagger}} \left\| \nabla_{\theta\theta}^2 \hat{Q}_T(\theta) - \mathcal{A}(\theta) \right\| \rightarrow_p 0 \quad (\text{TS.29})$$

---

<sup>2</sup>An equivalent synthetic formulation of (TS-A.iv) is that  $\{\ell_t(\theta)\}$  satisfies a uniform weak law of large numbers (UWLLN), see e.g. Wooldridge (1994).

where in this case  $\|\cdot\|$  is the matrix Euclidean norm and  $\mathcal{A}(\theta)$  is a (non-stochastic) matrix continuous at  $\theta^\dagger$  satisfying<sup>3</sup>

$$\begin{aligned}\mathcal{A}(\theta^\dagger) &:= \lim_{T \rightarrow \infty} E \left( \nabla_{\theta\theta}^2 \hat{Q}_T(\theta^\dagger) \right) = \lim_{T \rightarrow \infty} E \left( \frac{1}{T} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t(\theta^\dagger) \right) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E \left( \nabla_{\theta\theta}^2 \ell_t(\theta^\dagger) \right) = E \left( \nabla_{\theta\theta}^2 \ell_t(\theta^\dagger) \right);\end{aligned}$$

(TS-A.viii)  $\mathcal{A} = \mathcal{A}(\theta^\dagger)$  is nonsingular.

Then

$$T^{1/2}(\hat{\theta}_T - \theta^\dagger) \xrightarrow{d} N(0_{\dim(\theta) \times 1}, \mathcal{A}^{-1} \mathcal{B} \mathcal{A}^{-1}) \quad (\text{TS.30})$$

where  $\mathcal{A} = \mathcal{A}(\theta^\dagger) = (\mathcal{I}_{\theta^\dagger, \infty}^{2D})^{-1}$  and  $\mathcal{B} = \mathcal{B}(\theta^\dagger) = (\mathcal{I}_{\theta^\dagger, \infty}^{OP})^{-1}$  (see the definitions of  $\mathcal{I}_{\theta^\dagger, \infty}^{2D}$  and in  $\mathcal{I}_{\theta^\dagger, \infty}^{OP}$  in ACF). The proof of the standard asymptotic normality result may be found, among others, in e.g. Wooldridge (1994).

We recall that the convergence behind eq. (TS.30) stems from a conventional mean value expansion of  $\nabla_{\theta} \hat{Q}_T(\hat{\theta}_T)$  around  $\theta^\dagger$ , which leads to

$$\begin{aligned}T^{1/2}(\hat{\theta}_T - \theta^\dagger) &= - \left\{ \nabla_{\theta\theta}^2 \hat{Q}_T(\bar{\theta}_T) \right\}^{-1} T^{1/2} \nabla_{\theta} \hat{Q}_T(\theta^\dagger) \\ &= - \left\{ T^{-1} \sum_{t=1}^T \nabla_{\theta\theta}^2 \ell_t(\bar{\theta}_T) \right\}^{-1} \left\{ T^{-1/2} \sum_{t=1}^T \nabla_{\theta} \ell_t(\theta^\dagger) \right\}\end{aligned} \quad (\text{TS.31})$$

where  $\bar{\theta}_T$  is on the segment connecting the points  $\hat{\theta}_T$  and  $\theta^\dagger$  in  $\mathcal{P}^D$ . By (TS-A.vii), the first term in eq. (TS.31) converges uniformly in probability to  $\mathcal{A}(\theta^\dagger)^{-1}$  (as  $\bar{\theta}_T \rightarrow_p \theta^\dagger$ ), while by (TS-A.vi) the second term converges in distribution to  $N(0_{\dim(\theta) \times 1}, \mathcal{B})$ .

Condition (TS-A.v) might be violated in some estimated DSGE models, because it is often found that the QML estimator  $\hat{\theta}_T$  tends to locate on the boundary of the optimization set. Obviously, when  $\theta^\dagger$  lies exactly on the boundary of  $\mathcal{P}^D$ , asymptotic normality does no longer hold.<sup>4</sup> It is the uniform convergence condition (TS-A.vii) which is key to strong/weak identification. As detailed in ACF, in weakly identified DSGE models, the uniform convergence of the Hessian to a non-stochastic matrix in eq. (TS.29) is no longer guaranteed, hence the expansion in eq. (TS.31) cannot be used for the QML estimator, see Andrews and Cheng (2012) and Andrews and Mikusheva (2015). Finally, the condition (TS-A.viii) is violated in unidentified DSGE models, i.e. when the rank condition in Assumption 3 of ACF fails.

## TS.5 FURTHER MONTE CARLO RESULTS

In this section we report additional Monte Carlo results other the ones in Section 5 of ACF. Section TS.5.1 deals with the ARMA(1,1) model and Section TS.5.2 investigates Andrews and Mikusheva's (2015) DSGE model. Section TS.5.3 completes the Monte Carlo experiment in ACF considering also the case  $T = 250$ .

<sup>3</sup>An equivalent synthetic formulation of condition (TS-A.vii) is that  $\{\nabla_{\theta\theta}^2 \ell_t(\theta)\}$  satisfies a UWLLN.

<sup>4</sup>See Morris (2016) for a discussion of the 'pileup' phenomenon in DSGE models.

### TS.5.1 THE ARMA(1,1) MODEL

The ARMA(1,1) model represents an interesting case study for our bootstrap approach as it is particularly suited to characterize strong and weak identification.

Let  $y_t$  be a scalar that obeys the ARMA(1,1) model:

$$y_t = (\pi + \beta)y_{t-1} + \omega_t - \pi\omega_{t-1} \quad , \quad \omega_t \sim iidN(0, 1), \quad t = 1, \dots, T \quad (\text{TS.32})$$

where  $y_0$  and  $\omega_0$  are given, and the vector of parameters before any restriction is imposed is  $\phi := (\phi_1, \phi_2)' := (\pi, \pi + \beta)'$ , where  $\beta$  can be interpreted as the difference between the autoregressive ( $\phi_2 = \pi + \beta$ ) and moving average ( $\phi_1 = \pi$ ) coefficients. It is seen that in the special case  $\beta = 0$  ( $\phi_2 = \phi_1$ ), the model collapses to

$$y_t = \omega_t$$

and the moving average parameter  $\pi$  is not identified. In this case, Assumptions 2-3 of ACF are violated and the conditions for the consistency of the QML estimator are not satisfied. When  $\beta$  is close to zero (but different from zero so that Assumptions 2-3 still hold), the likelihood function of the ARMA(1,1) model is relatively flat in the direction of  $\pi$ . This model satisfies the parameterization of Andrews and Cheng (2012) and their asymptotic theory can be applied. Andrews and Mikusheva (2015) show that in the weak identification case, if the parameter  $\beta$  is defined through the embedding  $\beta_T = C/T^{1/2}$  for some constant  $C$ , then suitably normalized versions of the measures of information  $\mathcal{I}_{\theta_0, T}^{OP}$  and  $\mathcal{I}_{\theta_0, T}^{2D}$  (see Section 3.2 of ACF) converge in the limit to different quantities, in particular:

$$D_T(\mathcal{I}_{\theta_0, T}^{OP} - \mathcal{I}_{\theta_0, T}^{2D})D_T' \rightarrow_p M_{\dim(\theta) \times \dim(\theta)},$$

where  $M_{\dim(\theta) \times \dim(\theta)}$  is a random matrix and  $D_T$  is a normalization (typically diagonal) matrix. The equation above implies that  $\mathcal{I}_{\theta_0, T}^{OP}$  and  $\mathcal{I}_{\theta_0, T}^{2D}$  are no longer interchangeable measures of information even if White's (1982) information matrix equality  $E(\check{\mathcal{I}}_{\theta_0, T}^{OP}(\theta_0) - \check{\mathcal{I}}_{\theta_0, T}^{2D}(\theta_0)) = 0_{\dim(\theta) \times \dim(\theta)}$  is valid.

We assume that the model in eq. (TS.32) is stationary and invertible and consider the testing problem

$$H_0'' : \pi = 0.4 \quad \text{against} \quad H_1'' : \pi \neq 0.4. \quad (\text{TS.33})$$

Under  $H_0''$ , the (minimal) state-space representation associated with eq. (TS.32) is given by

$$Z_{m,t} = \begin{pmatrix} 0.4 + \theta & 1 \\ 0 & 0 \end{pmatrix} Z_{m,t-1} + \begin{pmatrix} 1 \\ -0.4 \end{pmatrix} \omega_t \quad (\text{TS.34})$$

$A(\phi_\theta)$   $B(\phi_\theta)$

$$y_t = \begin{pmatrix} 1, 0 \end{pmatrix} Z_{m,t} ; \quad (\text{TS.35})$$

$H^+(\phi_\theta)$

hence,  $\theta = \beta$  is the structural parameters that appears under the null. The associated (minimal) state-space representation without imposing the restriction (i.e. under  $H_1''$ ) is given by

$$Z_{m,t} = \begin{pmatrix} \pi + \beta & 1 \\ 0 & 0 \end{pmatrix} Z_{m,t-1} + \begin{pmatrix} 1 \\ -\pi \end{pmatrix} \omega_t \quad (\text{TS.36})$$

$A(\phi)$   $B(\phi)$



$$y_t = \underset{H^+(\phi)}{(1, 0)} Z_{m,t} \quad (\text{TS.37})$$

where the parameters  $\phi_1 = \pi$  and  $\phi_2 = (\pi + \beta)$  satisfy the conditions  $-1 < \phi_1 < 1$ ,  $-1 < \phi_2 < 1$ , which ensure stationarity and invertibility. In this specific example, the relationship  $\phi_\theta = g(\theta)$  implied by the null hypothesis  $H_0''$  is linear and can be summarized by the expression:

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \pi \\ \pi + \beta \end{pmatrix} = \text{under } H_0'' = \begin{pmatrix} 0.4 \\ 0.4 + \theta \end{pmatrix} = \phi_\theta.$$

We generate  $M = 1000$  samples of length  $T = 100$  and  $T = 250$  from the ARMA (1,1) model under the null  $H_0''$  in eq. (TS.33), setting  $y_0 = \omega_0 = 0$  and postulating a Gaussian distribution for  $\omega_t$ . For each replication, samples of  $T + 200$  observations are actually generated, with the first 200 observations being discarded. We consider different scenarios depending on the values assumed by the parameter  $\theta (= \beta)$ . As in Andrews and Cheng (2012), we select  $\theta (= \beta)$  from the set  $\{-0.76, -0.05\}$ , hence our results are partly comparable with their work. The data generating process which corresponds to the case of strong identification is obtained with  $\beta (= \theta) = -0.76$ , while  $\beta (= \theta) = -0.05$  characterizes the weak identification scenario.

For each data generating process considered and on each generated sample, we test the hypothesis  $H_0''$  against  $H_1''$  in (TS.33) by using a QLR test and compute the bootstrap p-value associated with the QLR test by applying the algorithm discussed in Section 4 of ACF. We use  $B = 99$  bootstrap replications for  $T = 100$  and  $B = 249$  bootstrap replications for  $T = 250$ . We run both a restricted bootstrap procedure (i.e. the bootstrap samples are generated imposing the null  $H_0''$ ) and an unrestricted bootstrap procedure (i.e. the bootstrap samples are generated without imposing the null). Both under  $H_0''$  and  $H_1''$ , the log-likelihood is maximized by the BFGS method, imposing that the optimization parameter spaces for the MA and AR coefficients are constrained to  $[-0.85, 0.85]$  and  $[-0.90, 0.90]$ , respectively. Since for this experiment the specified likelihood is correctly Gaussian, throughout this section we use the terms ‘ML estimator’ and ‘LR test’ in place of ‘QML estimator’ and ‘QLR test’. To evaluate the empirical size of the normality tests for strong/weak identification, we apply our bootstrap algorithm for different choices of  $\tilde{B}$  (out of  $B$ ) in the set  $\{10, 20, 30, 40, 99\}$  when we consider the case  $T = 100$ , and in the set  $\{20, 30, 40, 249\}$  when we consider the case  $T = 250$ . The so-obtained bootstrap realizations  $\hat{\theta}_{T:1}^*, \dots, \hat{\theta}_{T:\tilde{B}}^*$  (under  $H_0''$ ) and  $\hat{\pi}_{T:1}^*, \dots, \hat{\pi}_{T:\tilde{B}}^*$  and  $\hat{\beta}_{T:1}^*, \dots, \hat{\beta}_{T:\tilde{B}}^*$  (under  $H_1''$ ) are used to compute Jarque and Bera’s (1997) (henceforth JB) and Shapiro and Wilk’s (1965) (henceforth SW) normality tests at the 5% nominal significance level.

#### STRONG IDENTIFICATION

We start our investigation from the case  $\beta_0 (= \theta_0) = -0.76$  which characterizes a strongly identified ARMA(1,1) process. Estimation and testing results are summarized in Table TS1. We notice that the ML estimates of the parameters  $\pi$  and  $\beta$  under  $H_1''$  are substantially similar to their (restricted) bootstrap counterparts and tend to converge to their true population values as  $T$  increases. For both  $T=100$  and  $T=250$ ,

the analytic (Hessian-based) standard errors associated with  $\hat{\pi}_T$  and  $\hat{\beta}_T$  are similar to the bootstrap standard errors. The similarity between sample estimates and bootstrap estimates is also observed for  $\hat{\beta}_T(=\hat{\theta}_T)$  when the null  $H_0''$  is imposed. For both  $T=100$  and  $T=250$ , the rejection frequency of the LR test for  $H_0''$  against  $H_1''$  is close to the nominal 5% level. The bootstrap counterpart of the LR test displays similar rejection frequencies. In particular, the bootstrap distribution of the LR test seems to match perfectly the  $\chi_1^2$  approximation (see also the discussion below).

As concerns our bootstrap misspecification approach, we apply univariate tests of normality on  $\tilde{B}$  (selected out of  $B$ ) realizations of the bootstrap structural parameter estimators. The associated empirical rejection frequencies are reported in Tables TS2a-TS2b. Specifically, Tables TS2a summarizes the results obtained with the restricted bootstrap, while TS2b summarizes the results obtained with the unrestricted bootstrap. Recall that from the analysis in Section 4.3 of ACF we need  $\tilde{B} = o(T)$  for the test to be asymptotically valid. This fact seems to be fully reflected in the simulation results, which show that the lower  $\tilde{B}$  relative to  $T$ , the closer is the empirical size to the 5% nominal significance level. Figure TS1 compares, for the case  $T = 250$ , the finite sample densities of the ML estimators of the parameters under  $H_0''$  and  $H_1''$  (left-panel) with the finite sample densities of the corresponding bootstrap estimators (right-panel), and contrasts these densities with the Gaussian. The finite sample densities in the left-panel are computed across the  $M = 1000$  Monte Carlo simulations, while the finite sample densities on the right-panel are computed across  $B = 249$  repetitions obtained in one of the Monte Carlo simulations. Overall, the results in Table TS1 and Tables TS2a-TS2b show that in strong identified models, the bootstrap works in the expected direction. In particular, our simple test for strong identification displays a good finite sample size coverage for proper choices of  $\tilde{B}$  relative to  $T$ .

#### WEAK IDENTIFICATION

We move to the case ( $\beta_0(= \theta_0) = -0.05$ ), which characterizes a weakly identified ‘with near cancelling roots’ ARMA(1,1) process. Estimation and testing results are summarized in Table TS3. In this case, the ML estimator of  $\pi$  is not consistent (Andrews and Cheng 2012) and this is fully reflected in the output of Table TS3. However, notice that under  $H_0''$  the parameter  $\pi$  is fixed, while  $\beta(=\theta)$  plays the role of nuisance parameter in the testing problem of  $H_0''$  against  $H_1''$ . Andrews and Mikusheva (2015) show that  $\hat{\beta}_T(=\hat{\theta}_T)$  is consistent and asymptotically normal under the null, and this can be noticed from our results. For both  $T=100$  and  $T=250$ , we also observe a substantial mismatch between the ML estimates of  $\pi$  and  $\beta$  and the average of the bootstrap replicates under  $H_1''$ . The mismatch between analytic (Hessian-based) and bootstrap standard errors seems to increase with  $T$ . Instead, the ML estimates of  $\beta(=\theta)$  under  $H_0''$  are substantially similar to the average of the bootstrap ML estimates and tend to converge to the true population values as  $T$  increases, consistently with Andrews and Mikusheva’s (2015) findings. Interestingly, for both  $T=100$  and  $T=250$ , the rejection frequency of the LR test for  $H_0''$  against  $H_1''$  and of its bootstrap analog is close to the nominal 5% level, suggesting that in this case the  $\chi_1^2$  distribution still represents a good approximation, as also confirmed by Andrews and Cheng (2012).

The empirical rejection frequencies of the normality tests associated with the bootstrap repetitions of the structural parameter estimators are reported in Tables TS4a-TS4b for different values of  $\tilde{B}$ . Table TS4a summarizes the results obtained with the restricted bootstrap and Table TS4b summarizes the results obtained with the unrestricted bootstrap. Normality tests are computed using the 5% nominal level of significance. Notably, in this scenario, our approach does not detect strong deviations from normality for  $\hat{\beta}_T^*(=\hat{\theta}_T^*)$  under  $H_0''$ , in line with Andrews and Mikusheva's (2015) result. Conversely, and as expected, the rejection frequencies of the normality test of  $\hat{\pi}_T^*$  under  $H_1''$  signal identification issues. For values of  $\tilde{B}$  for which we have a satisfactorily size control in the case of strong identification ( $\tilde{B} = 20$ ,  $\tilde{B} = 30$ ), the rejection frequency of the test of normality of  $\hat{\pi}_T^*$  under  $H_1''$  lies around 15% and 50%. If we combine these evidences with the discrepancy observed between analytic standard errors and bootstrap standard errors, we can conclude that our approach detects the failure of the regularity conditions for standard inference rather well. The bootstrap seems to be informative and useful for practitioners.

### TS.5.2 ANDREWS AND MIKUSHEVA'S (2015) DSGE MODEL

We consider the DSGE model investigated in Andrews and Mikusheva (2015), given by the system of equations:

$$x_t = E_t x_{t+1} - (r_t - E_t \pi_{t+1} - \rho \Delta a_t), \quad (\text{TS.38})$$

$$\pi_t = b E_t \pi_{t+1} + \kappa x_t + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2) \quad (\text{TS.39})$$

$$r_t = \lambda r_{t-1} + (1 - \lambda) \phi_\pi \pi_t + (1 - \lambda) \phi_x x_t + u_t \quad (\text{TS.40})$$

$$\Delta a_t = \rho \Delta a_{t-1} + \varepsilon_{a,t}, \quad \varepsilon_{a,t} \sim \text{WN}(0, \sigma_a^2) \quad (\text{TS.41})$$

$$u_t = \delta u_{t-1} + \varepsilon_{u,t}, \quad \varepsilon_{u,t} \sim \text{WN}(0, \sigma_u^2) \quad (\text{TS.42})$$

which comprise a linearized output Euler equation (eq. (TS.38)) with autoregressive exogenous shocks  $\Delta a_t$  (eq. (TS.41)), a Phillips curve (eq. (TS.39)) with exogenous White Noise shock  $\varepsilon_t$ , and a monetary policy rule (eq. (TS.40)) with autoregressive exogenous shock  $u_t$  (eq. (TS.42)). Andrews and Mikusheva (2015) show that under the simplifying assumption  $\lambda = 0$ ,  $\phi_x = 0$ ,  $\phi_\pi = b^{-1}$  and  $\sigma^2 = 0$ , the sixth-dimensional vector  $\theta := (b, \kappa, \rho, \delta, \sigma_u^2, \sigma_a^2)'$  is point-identified if  $0 < b < 1$ ,  $\kappa > 0$ ,  $\sigma_u^2 > 0$ ,  $\sigma_a^2 > 0$ ,  $0 < \delta < \rho < 1$ , which means that for all  $\theta$ s with elements satisfying  $0 < b < 1$ ,  $\kappa > 0$ ,  $\sigma_u^2 > 0$ ,  $\sigma_a^2 > 0$ ,  $0 < \delta < \rho < 1$ , our Assumptions 2-3 are satisfied. Interestingly, Andrews and Mikusheva (2015) show that identification fails when  $\rho = \delta$ ; more precisely, when  $\rho = \delta$ , the parameter  $\theta := (b, \kappa, \rho, \delta, \sigma_u^2, \sigma_a^2)'$  loses two degrees of identification, meaning that two elements among  $b, \kappa, \sigma_u^2, \sigma_a^2$  are not separately identifiable. Instead, when  $\rho - \delta$  is different from zero but close to zero, there is weak identification and statistical inference on the remaining parameters may become unreliable.

We consider two DGPs. In the former, denoted AM-DGP1, other than the simplifying assumptions  $\lambda = 0$ ,  $\phi_x = 0$ ,  $\phi_\pi = b^{-1}$  and  $\sigma^2 = 0$ , we add the restriction that  $b$  and  $\sigma_a^2$  are known and fixed at their DGP values. In this case, even when  $\rho = \delta$ , the vector  $\theta_1 := (\kappa, \rho, \delta, \sigma_u^2)'$  is point-identified (Assumptions 2-3 in ACF are valid). We expect that the conditions for standard inference on  $\theta_1$  are valid in model AM-DGP1. In the latter,

denoted AM-DGP2, we retain the simplifying assumptions  $\lambda = 0$ ,  $\phi_x = 0$ ,  $\phi_\pi = b^{-1}$  and  $\sigma^2 = 0$ , but relax the hypothesis that  $\sigma_a^2$  is known (still keeping the discount factor  $b$  fixed at its DGP value, 0.99). In this case, despite the vector  $\theta_2 := (\kappa, \rho, \delta, \sigma_u^2, \sigma_a^2)'$  is point-identified and Assumptions 2-3 in ACF are valid, the distance  $\rho - \delta$  may affect the validity of the standard regularity conditions for inference. The DGP values used for  $\theta_1$  and  $\theta_2$  are taken from Table 1 of Andrews and Mikusheva (2015).

For both AM-DGP1 and AM-DGP2, we generate  $M = 1000$  samples of length  $T=100$  and  $T=250$  from the minimal ABCD form associated with the DSGE model, assuming that  $(\varepsilon_{a,t}, \varepsilon_t, \varepsilon_{u,t})'$  is Gaussian with diagonal covariance matrix, fixing the initial conditions to zero. For each replication, a sample of  $T + 200$  observations is actually generated and the first 200 observations are then discarded. On each generated dataset, we apply the bootstrap procedure summarized in Section 4 of ACF to analyze: (i) the empirical size of the QLR test for the CER; (ii) the bootstrap-adjusted empirical size of the QLR test for the CER (using  $B=99$  bootstrap repetitions for the case  $T = 100$  and  $B=249$  bootstrap replications for the case  $T = 250$ ); (iii) the empirical rejection frequency of our bootstrap misspecification test (in particular JB normality test) using the 5% nominal significance level. Results relative to AM-DGP1 are summarized in Table TS5 and TS6. Results relative to AM-DGP2 are summarized in Table TS7 and TS8. Since the data are simulated assuming a Gaussian distribution for the shocks, throughout we use ‘ML’ and ‘LR’ in place of ‘QML’ and ‘QLR’, respectively.

#### AM-DGP1

Results are summarized in Table TS5 for  $T=100$  and in Table TS6 for  $T=250$ . The ML estimates of the structural parameters and analytic (Hessian-based) standard errors are substantially similar to the bootstrap ML estimates and bootstrap standard errors, as it is expected in strongly identified DSGE model. For  $T=100$ , using the asymptotic critical values (taken from the  $\chi_3^2$  distribution), the rejection frequency of the LR test for the CER is 4.8%, already close in this case to the 5% nominal significance level. Using the bootstrap procedure, this empirical size collapses to 2.7%. For  $T=250$ , the rejection frequency based on asymptotic critical value is 5.8% and becomes 4.7% (still close to the 5% significance level) with the bootstrap. The bootstrap seems to work in the right directions in the presence of strong identification.

In the bottom panel of TS5 and TS6 we report the result of our misspecification approach. We observe for  $T=100$  and  $T=250$ , ‘physiologic’ empirical rejection frequencies for the normality tests which fall in the range 4.5%-10% for small values of  $\tilde{B}$ , as opposed to the 5% nominal significance level. The rejection frequencies increase, as expected, when we use the procedure summarized in the Remark 4.7 of ACF. As  $\tilde{B}$  increases relative to  $T$ , therefore violating the condition  $\tilde{B}/T \rightarrow 0$ , the misspecification test over-rejects the null hypothesis. We can conclude that conditionally on the calibrated parameters, the estimated Andrews and Mikusheva’s (2015) DSGE model with parameters  $\theta_1 := (\kappa, \rho, \delta, \sigma_u^2)'$  seems to meet the conditions that ensure strong identification.

## AM-DGP2

Table TS7 and Table TS8 summarize a more involved situation compared to the results reported in Table TS5 and TS6. For both  $T=100$  and  $T=250$ , the ML estimates of the structural parameters and their bootstrap analogs have similar magnitudes. However we observe some discrepancies between analytic (Hessian-based) and bootstrap standard errors for some parameters (e.g.  $\sigma_a^2$ ), even when the sample size is increased. For  $T = 100$ , using the asymptotic critical values (taken from the  $\chi_2^2$  distribution), the rejection frequency of the LR test for the CER is 3.6% as opposed to the 5% nominal significance level, and the empirical size of the bootstrap version of the test is 4.1%. For  $T=250$ , we obtain 1.1% and 2.4%, respectively. The tendency of the LR test to under-reject the CER as the sample size increases indicates that there might be issues as concerns the regularity conditions for standard inference. The finite sample performance of the bootstrap, however, seems to provide reasonable results also in these cases. More precisely, the bootstrap still improves on the asymptotic LR test by bringing the empirical rejection frequency close to the chosen nominal level.

Focusing on our bootstrap misspecification approach, in the bottom panels of Table TS7 and Table TS8 the rejection frequencies of the normality tests are in the range 5.3%-30% for  $T=100$  ( $\tilde{B} = 10$ ) and in the range 4.5%-40% for  $T=250$  ( $\tilde{B} = 20$ ). In particular, the rejection frequency associated with the normality of  $\hat{\sigma}_{a,T}^{2,*}$  is 29.5% and 39.5%, respectively. Based on these evidences, we can conclude that conditionally on the calibrated parameters, Andrews and Mikusheva's (2015) DSGE model with parameters  $\theta_2 := (\kappa, \rho, \delta, \sigma_a^2, \sigma_u^2)'$  does not meet the regularity conditions that endure standard asymptotic inference. However, albeit some parameters are affected by weak identification issues, bootstrap-based finite sample QML inference seems to be still reliable in this model.

### TS.5.3 MORE ON THE AN AND SCHORFHEIDE'S (2007) DSGE MODEL

This section completes the results of the Monte Carlo experiment discussed in Section 5 of ACF. Table 1 in ACF refers to AS-DGP1 version of An and Schorfheide's (2007) and summarizes results obtained with  $T=100$ . Table TS9 in this technical supplement extends the analysis to the case  $T=250$ . Table 2 in ACF refers to AS-DGP2 version of An and Schorfheide's (2007) and summarizes results obtained with  $T=100$ . Table TS10 in this technical supplement extends results to the case  $T=250$ .

The results in Tables TS9-TS10 lead us to the same conclusions as in Section 5 of ACF.

## TS.6 FURTHER EMPIRICAL RESULTS

In this section, we focus on the empirical results summarized in Section 6 and Table 3 of ACF. It is instructive for our purposes to compare the results obtained for the models AS-M1 and AS-M2 using our bootstrap approach with what one would obtain using identification-robust methods.

Consider model AS-M1 first. The unknown parameters are in  $\theta_1 := (\kappa, \rho_r, \sigma_g^2)'$

( $\dim(\theta_1)=3$ ). Following Guerron-Quintana *et al.* (2013), we consider the ‘new’ testing problem

$$H'_0 : \phi = \phi_{\check{\theta}_1} = g(\check{\theta}_1) \text{ for some } \theta_1 = \check{\theta}_1 \in \mathcal{P}_1^D \text{ against } H'_1 : \phi \neq \phi_{\check{\theta}_1},$$

where  $\check{\theta}_1$  is a ‘guess’ about the true value of  $\theta_1$ , and  $\mathcal{P}_1^D$  is the subset of the determinacy region of the parameter space  $\mathcal{P}^D$  which involves the elements of  $\theta_1$ . The composite null hypothesis  $H'_0$  establishes that the CER are valid in correspondence of the (known) parameter point  $\theta_1 = \check{\theta}_1 \in \mathcal{P}_1^D$ . Thus, the evaluation of the log-likelihood of the DSGE model under  $H'_0$  does not require any estimation procedure. Under  $H'_0$ , the QLR test  $QLR_T(\check{\theta}_1) := -2[\ell_{o,T}(\phi_{\check{\theta}_1}) - \ell_{o,T}(\hat{\phi}_T)]$  is asymptotically pivotal and  $\chi_{\dim(\phi)}^2$ -distributed regardless of the strength of identification. The null  $H'_0$  is not rejected at the  $100\eta\%$  nominal significance level if  $QLR_T(\check{\theta}_1) < c_{\chi_{\dim(\phi)}^2}^\eta$ . Considering a set of possible choices  $\theta_1 = \check{\theta}_1 \in \mathcal{P}_1^D$ , the inversion of the test for  $H'_0$  against  $H'_1$  produces the identification-robust confidence set  $\mathcal{C}_{1-\eta}^{LR} := \{\check{\theta}_1 \in \mathcal{P}_1^D, QLR_T(\check{\theta}_1) < c_{\chi_{\dim(\phi)}^2}^\eta\}$ . The set  $\mathcal{C}_{1-\eta}^{LR}$  contains all parameters points in  $\mathcal{P}_1^D$  for which the CER are not rejected at the nominal level  $\eta$ , and has  $100(1 - \eta)\%$  asymptotic coverage. The CER implied by the DSGE model are rejected when  $\mathcal{C}_{1-\eta}^{LR}$  is empty and accepted otherwise. Obviously,  $\mathcal{C}_{1-\eta}^{LR}$  will be empty when  $QLR_T(\hat{\theta}_{1,T}) := \min_{\check{\theta}_1 \in \mathcal{P}_1^D} QLR_T(\check{\theta}_1) \geq c_{\chi_{\dim(\phi)}^2}^\eta$ .

From the upper panel of Table 3 we notice that as concerns model AS-M1, the QLR test for the CER, equal to 176.11, implies the rejection of the CER when the 5% asymptotic critical value from the  $\chi_{13}^2$  distribution (22.36) is employed, meaning that whatever method a practitioner uses to invert the test,  $\mathcal{C}_{0.95}^{LR}$  will be empty. Thus, while our bootstrap approach provides some empirical support to the estimated DSGE model, both ‘standard’ and identification-robust methods rejected model AS-M1 at the 5% nominal significance level.

Consider model AS-M2 now. The unknown parameters are in  $\theta_2 := (\kappa, \psi_1, \rho_r, \sigma_z^2, \sigma_g^2, \sigma_r^2)'$  ( $\dim(\theta_2)=6$ ). From the lower panel of Table 3 we note that the QLR test for the CER equals 83.25 and implies rejection of the CER using the 5% asymptotic critical value. That is, the 95% identification-robust confidence set  $\mathcal{C}_{0.95}^{LR}$  is empty. An empty identification-robust confidence set is also obtained at the 90% (nominal) asymptotic coverage ( $\mathcal{C}_{0.90}^{LR}$ ).

## REFERENCES

- An, S., Schorfheide, F. (2007), Bayesian analysis of DSGE models, *Econometric Reviews* 26, 113-172.
- Anderson, B.D.O. and Moore, J. (1979), *Optimal filtering*, Englewood Cliffs, NJ: Prentice-Hall.
- Andreasen, M. M. (2010), How to maximize the likelihood function for a DSGE model, *Computational Economics* 35, 127-154.
- Andrews, D.W.K., Cheng, X. (2012), Estimation and inference with weak, semi-strong, and strong identification, *Econometrica* 80, 2153-2211.

- Andrews, I., Mikusheva, A. (2015), Maximum likelihood inference in weakly identified dynamic stochastic general equilibrium models, *Quantitative Economics* 6, 123-152.
- Angelini, G., Cavaliere, G. and Fanelli, L. (2016), Bootstrapping DSGE models.
- Binder, M., Pesaran, M. H. (1995), Multivariate rational expectations models and macroeconomic modelling: A review and some new results, in M. Pesaran and M. Wickens (eds.) *Handbook of Applied Econometrics*, Oxford: Blackwell Publishing Ltd, 139-187.
- Fernández-Villaverde, J., Rubio-Ramírez, J., Sargent, T., Watson, M. (2007), ABCs (and Ds) of understanding VARs, *American Economic Review* 97, 1021-1026.
- Guerron-Quintana P, Inoue A, Kilian L. (2013), Frequentist inference in weakly identified DSGE models, *Quantitative Economics* 4, 197-229.
- Iskrev, N. (2008), Evaluating the information matrix in linearized DSGE models, *Economic Letters* 99, 607-610.
- Iskrev, N. (2010), Local identification in DSGE models, *Journal of Monetary Economics* 57, 189-202.
- Jarque, C. M., Bera, A. K. (1987), A test for normality of observations and regression residuals, *International Statistical Review* 55, 163-172.
- Komunjer, I., Ng, S. (2011), Dynamic identification of dynamic stochastic general equilibrium models, *Econometrica* 79, 1995-2032.
- Mavroeidis, S. (2005), Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve, *Journal of Money Credit and Banking* 37, 421-448.
- Morris, S. (2016), DSGE pileups, Working Paper, Department of Economics, Bowdoin College.
- Shapiro, S. S., Wilk, M. B. (1965), An analysis of variance test for normality (complete samples), *Biometrika* 52, 591-611.
- Qu, Z., Tkachenko, D. (2012), Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models, *Quantitative Economics* 3, 95-132.
- Uhlig, H. (1999), A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily, in R. Marimon and A. Scott (eds.), *Computational Methods for the Study of Dynamic Economies*, Oxford: Oxford University Press.
- Wooldridge, J.M. (1994), Estimation and inference for dependent processes, *Handbook of Econometrics*, R.F. Engle and D.L. McFadden (eds.), Vol. IV, Chap. 45.

TABLE TS1. MC experiment. Strong identification, ARMA(1,1) model. ML estimates and rejection frequencies of LR test.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.76$						
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$						
T=100	Monte Carlo			Bootstrap (B=99)		
	under $H_0''$	under $H_1''$		under $H_0''$	under $H_1''$	
	$\hat{\beta}_T(=\hat{\theta}_T)$	$\hat{\pi}_T$	$\hat{\beta}_T$	$\hat{\beta}_T^*(=\hat{\theta}_T^*)$	$\hat{\pi}_T^*$	$\hat{\beta}_T^*$
average	-0.7542	0.4089	-0.7597	-0.7495	0.4089	-0.7541
s.e.	0.0949	0.1487	0.1040	0.0924	0.1508	0.1012
Hessian-based s.e.	0.0948	0.1418	0.1022	0.0961	0.1586	0.1077
	$Rej(LR_T) = 0.0522$			$Rej(LR_T^*) = 0.0400$		
T=250	Monte Carlo			Bootstrap (B=249)		
	under $H_0''$	under $H_1''$		under $H_0''$	under $H_1''$	
	$\hat{\beta}_T(=\hat{\theta}_T)$	$\hat{\pi}_T$	$\hat{\beta}_T$	$\hat{\beta}_T^*(=\hat{\theta}_T^*)$	$\hat{\pi}_T^*$	$\hat{\beta}_T^*$
average	-0.7588	0.4028	-0.7602	-0.7574	0.4039	-0.7591
s.e.	0.0585	0.0893	0.0632	0.0588	0.0903	0.0632
Hessian-based s.e.	0.0594	0.0976	0.0674	0.0598	0.0984	0.0667
	$Rej(LR_T) = 0.0493$			$Rej(LR_T^*) = 0.0470$		

NOTES: Results are based on M=1000 Monte Carlo replications. ‘Average’ is the average of the ML estimator across Monte Carlo replications. ‘MC s.e.’ indicates the standard errors computed across Monte Carlo replications. ‘Hessian-based s.e.’ indicates the average across Monte Carlo replications of the analytic standard errors computed from the Hessian matrix. ‘Rej(·)’ denotes rejection frequency (across Monte Carlo simulations).  $LR_T$  is the LR test for  $H_0''$  vs  $H_1''$ .  $LR_T^*$  is the bootstrap analog of  $LR_T$  and is obtained with  $B$  replications.



TABLE TS2a. MC experiment. Strong identification, ARMA(1,1) model. Rejection frequencies of misspecification normality tests. Restricted bootstrap.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.76$					
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$					
T =100	$\tilde{B} = 10$	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 99$
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0520	0.0670	0.0780	0.0960	0.1400
$\hat{\pi}_T^*$ under $H_1''$	0.0800	0.1130	0.1280	0.1430	0.2310
$\hat{\beta}_T^*$ under $H_1''$	0.0690	0.0660	0.0600	0.680	0.0720
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0490	0.0460	0.0640	0.0810	0.1160
$\hat{\pi}_T^*$ under $H_1''$	0.0710	0.1040	0.1250	0.1370	0.2320
$\hat{\beta}_T^*$ under $H_1''$	0.0510	0.0540	0.0590	0.0690	0.0750
<hr/>					
T =250	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 249$	
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0670	0.0710	0.0790	0.1370	
$\hat{\pi}_T^*$ under $H_1''$	0.1020	0.0920	0.0860	0.1910	
$\hat{\beta}_T^*$ under $H_1''$	0.0630	0.0550	0.0660	0.0770	
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0510	0.0590	0.0620	0.1230	
$\hat{\pi}_T^*$ under $H_1''$	0.0770	0.0950	0.0930	0.2030	
$\hat{\beta}_T^*$ under $H_1''$	0.0480	0.0550	0.0520	0.0550	

NOTES: Results are based on M=1000 Monte Carlo replications. Rejection frequencies refer to Shapiro and Wilk's (1965) and Jarque and Bera's (1987) normality test. Normality tests are computed using the 5% nominal significance level.

TABLE TS2b. MC experiment. Strong identification, ARMA(1,1) model. Rejection frequencies of misspecification normality tests. Unrestricted bootstrap.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.76$					
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$					
T =100	$\tilde{B} = 10$	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 99$
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0480	0.0840	0.0740	0.1000	0.1360
$\hat{\pi}_T^*$ under $H_1''$	0.0860	0.1440	0.1780	0.2000	0.2880
$\hat{\beta}_T^*$ under $H_1''$	0.0420	0.0600	0.0720	0.1040	0.0960
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0480	0.0740	0.0700	0.0840	0.1240
$\hat{\pi}_T^*$ under $H_1''$	0.0820	0.1240	0.1440	0.1780	0.2880
$\hat{\beta}_T^*$ under $H_1''$	0.0420	0.0720	0.0680	0.0960	0.0740
T =250	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 249$	
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0620	0.0580	0.0620	0.1460	
$\hat{\pi}_T^*$ under $H_1''$	0.0900	0.0960	0.0940	0.1800	
$\hat{\beta}_T^*$ under $H_1''$	0.0500	0.0660	0.0560	0.0580	
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0480	0.0500	0.0560	0.1280	
$\hat{\pi}_T^*$ under $H_1''$	0.0740	0.0940	0.0900	0.2040	
$\hat{\beta}_T^*$ under $H_1''$	0.0460	0.0500	0.0440	0.0660	

NOTES: See Table TS2a.

TABLE TS3. MC experiment. Weak identification, ARMA(1,1) model. ML estimates and rejection frequencies of LR test.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.05$						
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$						
T=100	Monte Carlo			Bootstrap (B=99)		
	under $H_0''$	under $H_1''$		under $H_0''$	under $H_1''$	
	$\hat{\beta}_T (= \hat{\theta}_T)$	$\hat{\pi}_T$	$\hat{\beta}_T$	$\hat{\beta}_T^* (= \hat{\theta}_T^*)$	$\hat{\pi}_T^*$	$\hat{\beta}_T^*$
average	-0.0584	0.1856	-0.0626	-0.0666	0.2261	-0.0683
s.e.	0.0949	0.4865	0.1149	0.0927	0.4813	0.1085
Hessian-based s.e.	0.0950	0.4362	0.0939	0.0964	2.3251	0.2864
$Rej(LR_T) = 0.0453$			$Rej(LR_T^*) = 0.0410$			
T=250	Monte Carlo			Bootstrap (B=249)		
	under $H_0''$	under $H_1''$		under $H_0''$	under $H_1''$	
	$\hat{\beta}_T (= \hat{\theta}_T)$	$\hat{\pi}_T$	$\hat{\beta}_T$	$\hat{\beta}_T^* (= \hat{\theta}_T^*)$	$\hat{\pi}_T^*$	$\hat{\beta}_T^*$
average	-0.0537	0.2021	-0.0586	-0.0576	0.2291	-0.0598
s.e.	0.0593	0.4700	0.0672	0.0591	0.4699	0.0664
Hessian-based s.e.	0.0596	0.4362	0.0624	0.0600	2.1502	0.1558
$Rej(LR_T) = 0.0532$			$Rej(LR_T^*) = 0.0480$			

NOTES: See Table TS1.

TABLE TS4a. MC experiment. Weak identification, ARMA(1,1) model. Rejection frequencies of misspecification normality tests. Restricted bootstrap.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.05$					
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$					
T =100	$\tilde{B} = 10$	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 99$
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0580	0.0600	0.0680	0.0760	0.1150
$\hat{\pi}_T^*$ under $H_1''$	0.1880	0.4540	0.7070	0.8640	0.9980
$\hat{\beta}_T^*$ under $H_1''$	0.0850	0.1090	0.1440	0.1670	0.3180
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0530	0.0460	0.0580	0.0600	0.0970
$\hat{\pi}_T^*$ under $H_1''$	0.0720	0.1070	0.1500	0.2040	0.9530
$\hat{\beta}_T^*$ under $H_1''$	0.0640	0.0780	0.0810	0.0960	0.1760
T =250	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 249$	
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0720	0.0640	0.0790	0.1230	
$\hat{\pi}_T^*$ under $H_1''$	0.4250	0.6600	0.8340	1.0000	
$\hat{\beta}_T^*$ under $H_1''$	0.1100	0.1670	0.1810	0.6860	
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0470	0.0530	0.0570	0.0980	
$\hat{\pi}_T^*$ under $H_1''$	0.1310	0.1750	0.2330	1.0000	
$\hat{\beta}_T^*$ under $H_1''$	0.0780	0.0960	0.1120	0.4240	

NOTES: See Table TS2a.

TABLE TS4b. MC experiment. Weak identification, ARMA(1,1) model. Rejection frequencies of misspecification normality tests. Unrestricted bootstrap.

DGP: ARMA(1,1), $\pi_0 = 0.4, \beta_0 = -0.05$					
$H_0'' : \pi = 0.4$ vs $H_1'' : \pi \neq 0.4$					
T =100	$\tilde{B} = 10$	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 99$
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0700	0.0920	0.1180	0.1380	0.2540
$\hat{\pi}_T^*$ under $H_1''$	0.2940	0.5360	0.7160	0.8280	0.9820
$\hat{\beta}_T^*$ under $H_1''$	0.1100	0.1380	0.1940	0.2180	0.4520
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0420	0.0760	0.0980	0.1200	0.2240
$\hat{\pi}_T^*$ under $H_1''$	0.1700	0.1840	0.2380	0.3080	0.8860
$\hat{\beta}_T^*$ under $H_1''$	0.0760	0.1060	0.1300	0.1460	0.2800
T =250	$\tilde{B} = 20$	$\tilde{B} = 30$	$\tilde{B} = 40$	$\tilde{B} = 249$	
Shapiro-Wilk					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0760	0.1000	0.0980	0.2980	
$\hat{\pi}_T^*$ under $H_1''$	0.5220	0.7080	0.8180	0.9900	
$\hat{\beta}_T^*$ under $H_1''$	0.1180	0.1820	0.2060	0.7540	
Jarque-Bera					
$\hat{\beta}_T^*(= \hat{\theta}_T^*)$ under $H_0''$	0.0700	0.0620	0.0820	0.2760	
$\hat{\pi}_T^*$ under $H_1''$	0.1900	0.2420	0.3200	0.9780	
$\hat{\beta}_T^*$ under $H_1''$	0.0860	0.1080	0.1260	0.5020	

NOTES: See Table TS2a.

TABLE TS5. MC experiment. AM-DGP1

$b_0$	$\kappa_0$	$\lambda_0$	$\phi_{x,0}$	$\phi_{\pi,0}$	$\rho_0$	$\delta_0$	$\sigma_0^2$	$\sigma_{u,0}^2$	$\sigma_{a,0}^2$	
0.99	0.10	0	0	$\frac{1}{b_0}$	0.85	0.35	0	0.20	0.30	
T=100		Monte Carlo					Bootstrap (B=99)			
			$\hat{\theta}_{1,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{1,T}^*$	s.e.		
	$\kappa$		0.1008	0.0068	0.0068		0.1022	0.0074		
	$\rho$		0.8488	0.0097	0.0104		0.8470	0.0100		
	$\delta$		0.3337	0.0945	0.0990		0.3195	0.0950		
	$\sigma_u^2$		0.2049	0.0443	0.0486		0.2088	0.0472		
			$Rej(LR_T) = 0.0479$				$Rej(LR_T^*) = 0.0272$			
Normality, JB		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.1067	0.1548	0.1987	0.2594	0.4895				
	$\hat{\rho}_T^*$	0.0837	0.1234	0.1548	0.1967	0.3138				
	$\hat{\delta}_T^*$	0.0544	0.0544	0.0565	0.0649	0.0941				
	$\hat{\sigma}_{u,T}^{2,*}$	0.0795	0.1088	0.1569	0.1946	0.3515				
Normality, JB (Proc.)		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.1799	0.3326	0.3975	0.4519	0.6276				
	$\hat{\rho}_T^*$	0.1464	0.2615	0.3033	0.3661	0.4916				
	$\hat{\delta}_T^*$	0.0481	0.0962	0.1213	0.1695	0.3222				
	$\hat{\sigma}_{u,T}^{2,*}$	0.1695	0.2741	0.3013	0.3703	0.5146				

NOTES: Results are based on M=1000 Monte Carlo replications. Upper panel: ' $\hat{\theta}_{1,T}$ ' is the average of the ML estimator across Monte Carlo replications; 's.e.' indicates the standard errors computed across Monte Carlo replications; 's.e.(Hess.)' indicates the average of the Hessian-based standard errors across Monte Carlo replications; ' $\hat{\theta}_{1,T}^*$ ' is the average across Monte Carlo repetitions of the bootstrap ML estimators obtained as average of the bootstrap replicates  $\hat{\theta}_{1,T:1}^*, \dots, \hat{\theta}_{1,T:B}^*$ ; the log-likelihood maximization under both  $H_0$  (with the CER) and  $H_1$  (without the CER) is obtained by the combining the Kalman filter with the 'CMA-ES' algorithm; 'Rej( $\cdot$ )' denotes rejection frequency across Monte Carlo replications,  $LR_T$  is the LR test for  $H_0$  against  $H_1$  and is computed using the 5% nominal significance level;  $LR_T^*$  is the bootstrap analog of  $LR_T$  based on  $B$  bootstrap repetitions. Lower panel: 'Normality, JB' reports rejection frequencies associated with Jarque and Bera's (1987) normality tests of the  $\tilde{B}$  bootstrap repetitions of the structural parameter estimators; 'Normality, JB (Proc.)' reports rejection frequencies associated with the procedure described in Remark 4.7 of ACF; the  $N = \text{int}(B/\tilde{B})$  sub-samples used for the procedure in Remark 4.7 are not overlapping; all normality tests are computed using the 5% nominal significance level.

TABLE TS6. MC experiment. AM-DGP1

$b_0$	$\kappa_0$	$\lambda_0$	$\phi_{x,0}$	$\phi_{\pi,0}$	$\rho_0$	$\delta_0$	$\sigma_0^2$	$\sigma_{u,0}^2$	$\sigma_{a,0}^2$	
0.99	0.10	0	0	$\frac{1}{b_0}$	0.85	0.35	0	0.20	0.30	
T=250		Monte Carlo					Bootstrap (B=249)			
			$\hat{\theta}_{1,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{1,T}^*$	s.e.		
	$\kappa$		0.1004	0.0042	0.0042		0.1010	0.0043		
	$\rho$		0.8494	0.0060	0.0060		0.8487	0.0061		
	$\delta$		0.3455	0.0597	0.0621		0.3395	0.0594		
	$\sigma_u^2$		0.2006	0.0284	0.0288		0.2020	0.0278		
			$Rej(LR_T) = 0.0583$				$Rej(LR_T^*) = 0.0471$			
Normality, JB			$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
	$\hat{\kappa}_T^*$		0.0724	0.0881	0.1115	0.3503				
	$\hat{\rho}_T^*$		0.0568	0.0724	0.0920	0.2270				
	$\hat{\delta}_T^*$		0.0450	0.0665	0.0705	0.1233				
	$\hat{\sigma}_{u,T}^{2,*}$		0.0822	0.0959	0.1037	0.3875				
Normality, JB (Bonf.)			$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
	$\hat{\kappa}_T^*$		0.1057	0.1879	0.2074	0.5988				
	$\hat{\rho}_T^*$		0.0802	0.1389	0.1683	0.4951				
	$\hat{\delta}_T^*$		0.0802	0.0724	0.1057	0.4344				
	$\hat{\sigma}_{u,T}^{2,*}$		0.1468	0.1703	0.1996	0.6125				

NOTES: See Table TS5.

TABLE TS7. MC experiment. AM-DGP2

$b_0$	$\kappa_0$	$\lambda_0$	$\phi_{x,0}$	$\phi_{\pi,0}$	$\rho_0$	$\delta_0$	$\sigma_0^2$	$\sigma_{u,0}^2$	$\sigma_{a,0}^2$	
0.99	0.10	0	0	$\frac{1}{b_0}$	0.40	0.35	0	0.20	0.30	
T=100		Monte Carlo					Bootstrap (B=99)			
			$\hat{\theta}_{2,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{2,T}^*$	s.e.		
	$\kappa$		0.1019	0.0112	0.0168		0.1036	0.0113		
	$\rho$		0.4120	0.0653	0.1024		0.4115	0.0648		
	$\delta$		0.3216	0.0706	0.1228		0.3103	0.0756		
	$\sigma_a^2$		0.4656	0.4199	0.4748		0.4043	0.3575		
	$\sigma_u^2$		0.1887	0.1180	0.1632		0.2081	0.1086		
			$Rej(LR_T) = 0.0363$				$Rej(LR_T^*) = 0.0405$			
Normality, JB		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.0531	0.0922	0.0852	0.0922	0.1746				
	$\hat{\rho}_T^*$	0.0768	0.0894	0.1075	0.1299	0.2709				
	$\hat{\delta}_T^*$	0.0712	0.0936	0.1201	0.1187	0.2137				
	$\hat{\sigma}_{a,T}^{2,*}$	0.2947	0.3534	0.5209	0.8156	0.9916				
	$\hat{\sigma}_{u,T}^{2,*}$	0.0880	0.1034	0.1159	0.1453	0.2989				
Normality, JB (Bonf.)		$\tilde{B}=10$	$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=99$				
	$\hat{\kappa}_T^*$	0.0587	0.1313	0.1662	0.2304	0.3743				
	$\hat{\rho}_T^*$	0.0894	0.1676	0.2151	0.2709	0.4525				
	$\hat{\delta}_T^*$	0.0824	0.1676	0.1997	0.2737	0.4399				
	$\hat{\sigma}_{a,T}^{2,*}$	0.4609	0.4721	0.4791	0.5419	0.9916				
	$\hat{\sigma}_{u,T}^{2,*}$	0.1229	0.1662	0.2039	0.2556	0.4637				

NOTES: See Table TS5 and replace  $\hat{\theta}_{1,T}$  ( $\hat{\theta}_{1,T}^*$ ) with  $\hat{\theta}_{2,T}$  ( $\hat{\theta}_{2,T}^*$ )



TABLE TS8. MC experiment. AM-DGP2

$b_0$	$\kappa_0$	$\lambda_0$	$\phi_{x,0}$	$\phi_{\pi,0}$	$\rho_0$	$\delta_0$	$\sigma_0^2$	$\sigma_{u,0}^2$	$\sigma_{a,0}^2$	
0.99	0.10	0	0	$\frac{1}{b_0}$	0.40	0.35	0	0.20	0.30	
T=250		Monte Carlo					Bootstrap (B=249)			
			$\hat{\theta}_{2,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{2,T}^*$	s.e.		
	$\kappa$		0.1004	0.0070	0.0075		0.1012	0.0075		
	$\rho$		0.4134	0.0469	0.0494		0.4168	0.0446		
	$\delta$		0.3377	0.0429	0.0514		0.3300	0.0483		
	$\sigma_a^2$		0.4312	0.3911	0.2727		0.4111	0.3273		
	$\sigma_u^2$		0.1818	0.0903	0.0874		0.1888	0.0848		
			$Rej(LR_T) = 0.0105$				$Rej(LR_T^*) = 0.0245$			
Normality, JB			$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
	$\hat{\kappa}_T^*$		0.0455	0.0490	0.0699	0.2203				
	$\hat{\rho}_T^*$		0.0839	0.1294	0.1399	0.4580				
	$\hat{\delta}_T^*$		0.0909	0.1084	0.1364	0.3881				
	$\hat{\sigma}_{a,T}^{2,*}$		0.3951	0.5175	0.8287	0.9930				
	$\hat{\sigma}_{u,T}^{2,*}$		0.1573	0.1888	0.2238	0.8042				
Normality, JB (Bonf.)			$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
	$\hat{\kappa}_T^*$		0.0804	0.0979	0.1294	0.4371				
	$\hat{\rho}_T^*$		0.1503	0.1818	0.2378	0.6224				
	$\hat{\delta}_T^*$		0.1958	0.1993	0.2832	0.6434				
	$\hat{\sigma}_{a,T}^{2,*}$		0.4266	0.4196	0.4336	0.9930				
	$\hat{\sigma}_{u,T}^{2,*}$		0.1783	0.2063	0.2552	0.8287				

NOTES: See Table TS5 and replace  $\hat{\theta}_{1,T}$  ( $\hat{\theta}_{1,T}^*$ ) with  $\hat{\theta}_{2,T}$  ( $\hat{\theta}_{2,T}^*$ ).

TABLE TS9. MC experiment. AS-DGP1

$\tau_0$	$\kappa_0$	$\psi_{1,0}$	$\psi_{2,0}$	$\beta_0$	$\rho_{r,0}$	$\rho_{g,0}$	$\rho_{z,0}$	$\sigma_{r,0}^2$	$\sigma_{g,0}^2$	$\sigma_{z,0}^2$	
2	0.33	1.5	0.125	0.9975	0.75	0.95	0.90	0.04	0.36	0.09	
T=250				Monte Carlo				Bootstrap (B=249)			
				$\hat{\theta}_{1,T}$	s.e.	s.e.(Hess.)	$\hat{\theta}_{1,T}^*$				s.e.
$\kappa$				0.3318	0.0244	0.0250	0.3329				0.0257
$\rho_r$				0.7500	0.0158	0.0110	0.7490				0.0116
$\sigma_g^2$				0.3562	0.0328	0.0319	0.3551				0.0308
				$Rej(LR_T) = 0.0660$				$Rej(LR_T^*) = 0.0320$			
Normality, JB				$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
$\hat{\kappa}_T^*$				0.0880	0.0880	0.1120	0.1840				
$\hat{\rho}_{r,T}^*$				0.0480	0.0880	0.0880	0.0720				
$\hat{\sigma}_{g,T}^{2,*}$				0.0640	0.0480	0.0560	0.0960				
Normality, JB (Proc.)				$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$				
$\hat{\kappa}_T^*$				0.0880	0.1360	0.1520	0.1840				
$\hat{\rho}_{r,T}^*$				0.0240	0.0400	0.0480	0.0720				
$\hat{\sigma}_{g,T}^{2,*}$				0.0640	0.0320	0.0400	0.0960				

NOTES: Results are based on M=500 Monte Carlo replications. Upper panel: ' $\hat{\theta}_{1,T}$ ' is the average of the ML estimator across Monte Carlo replications; 's.e.' indicates the standard errors computed across Monte Carlo replications; 's.e.(Hess.)' indicates the average of the Hessian-based standard errors across Monte Carlo replications; ' $\hat{\theta}_{1,T}^*$ ' is the average across Monte Carlo repetitions of the bootstrap ML estimators obtained as average of  $\hat{\theta}_{1,T:1}^*, \dots, \hat{\theta}_{1,T:B}^*$ ; the log-likelihood maximization under both  $H_0$  (with the CER) and  $H_1$  (without the CER) is obtained by the combining the Kalman filter with the 'CMA-ES' algorithm; 'Rej( $\cdot$ )' denotes rejection frequency across Monte Carlo replications,  $LR_T$  is the LR test for  $H_0$  against  $H_1$  and is computed using the 5% nominal significance level and  $LR_T^*$  is the bootstrap analog of  $LR_T$  based on  $B$  bootstrap repetitions. Lower panel: 'Normality, JB' reports rejection frequencies associated with Jarque and Bera's (1987) normality tests of the  $\tilde{B}$  bootstrap repetitions of the structural parameter estimators; 'Normality, JB' (Proc.) reports rejection frequencies associated with the procedure described in Remark 4.7; the  $N = \text{int}(B/\tilde{B})$  sub-samples used for the procedure in Remark 4.7 are not overlapping; all normality tests are computed using the 5% nominal significance level.

TABLE TS10. MC experiment. AS-DGP2

$\tau_0$	$\kappa_0$	$\psi_{1,0}$	$\psi_{2,0}$	$\beta_0$	$\rho_{r,0}$	$\rho_{g,0}$	$\rho_{z,0}$	$\sigma_{r,0}^2$	$\sigma_{g,0}^2$	$\sigma_{z,0}^2$
2	0.33	1.5	0.125	0.9975	0.75	0.95	0.90	0.04	0.36	0.09
T=250				Monte Carlo			Bootstrap (B=249)			
			$\hat{\theta}_{2,T}$	s.e.	s.e.(Hess.)		$\hat{\theta}_{2,T}^*$	s.e.		
	$\kappa$		0.3338	0.0367	0.0363		0.3419	0.0389		
	$\psi_1$		1.5049	0.0982	0.0745		1.5187	0.1067		
	$\rho_r$		0.7500	0.0158	0.0110		0.7500	0.0167		
	$\sigma_r^2$		0.0398	0.0038	0.0037		0.0385	0.0038		
	$\sigma_g^2$		0.3562	0.0328	0.0319		0.3530	0.0306		
	$\sigma_z^2$		0.0906	0.0143	0.0121		0.0898	0.0149		
			$Rej(LR_T) = 0.0640$				$Rej(LR_T^*) = 0.0403$			
Normality, JB				$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$			
	$\hat{\kappa}_T^*$			0.1007	0.1141	0.1208	0.5906			
	$\hat{\psi}_{1,T}^*$			0.1208	0.0872	0.1074	0.5168			
	$\hat{\rho}_{r,T}^*$			0.0403	0.0470	0.0336	0.1208			
	$\hat{\sigma}_{r,T}^{2,*}$			0.0336	0.0470	0.0604	0.2081			
	$\hat{\sigma}_{g,T}^{2,*}$			0.0604	0.0470	0.0336	0.1208			
	$\hat{\sigma}_{z,T}^{2,*}$			0.0872	0.0403	0.0604	0.2617			
Normality, JB (Proc.)				$\tilde{B}=20$	$\tilde{B}=30$	$\tilde{B}=40$	$\tilde{B}=249$			
	$\hat{\kappa}_T^*$			0.1074	0.1342	0.1544	0.5906			
	$\hat{\psi}_{1,T}^*$			0.2550	0.3087	0.3289	0.5168			
	$\hat{\rho}_{r,T}^*$			0.0537	0.0940	0.0671	0.1208			
	$\hat{\sigma}_{r,T}^{2,*}$			0.0805	0.1007	0.0604	0.2081			
	$\hat{\sigma}_{g,T}^{2,*}$			0.0604	0.0537	0.0470	0.1208			
	$\hat{\sigma}_{z,T}^{2,*}$			0.0671	0.0805	0.1074	0.2617			

NOTES: See Table TS9 and replace  $\hat{\theta}_{1,T}$  ( $\hat{\theta}_{1,T}^*$ ) with  $\hat{\theta}_{2,T}$  ( $\hat{\theta}_{2,T}^*$ ).